



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number:** 122947

**TO: Frank Prats**  
**Location: REM/3A41/3E71**  
**Art Unit: 1651**  
**Thursday, May 27, 2004**

**Case Serial Number: 09/807146**

**From: Mary Jane Ruhl**  
**Location: Biotech-Chem Library**  
**Remsen 1-B55**  
**Phone: 571-272-2524**

**maryjane.ruhl@uspto.gov**

### **Search Notes**

Examiner Prats,

Here are the results for your recent search request.

Please feel free to contact me if you have any questions about these results.

Thank you for using STIC services. We appreciate the opportunity to serve you.

Sincerely,

Mary Jane Ruhl  
Technical Information Specialist  
STIC  
CM-1, Rm. 6-A-06  
605-1155

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 18:51:36 ; Search time 62 Seconds  
(without alignments)  
2898.391 Million cell updates/sec

Title: US-09-807-146-1  
Perfect score: 3408  
Sequence: 1 MLPTQVGLILQYIKRIL.....VSLNODLTLQYQVWMLRFA 636

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3408	100.0	636	3 AAY81945	Aay81945 Neisseria
2	3408	100.0	636	3 AAY90979	Aay90979 N. polysa
3	3408	100.0	636	3 AAY87831	Aay87831 N. polysa
4	3360	98.6	856	4 AAB70883	Aab70883 Expressio
5	2540	74.5	614	2 AAR88386	Aar88386 Neisseria
6	2540	74.5	614	2 AAR88633	Aar88633 Neisseria
7	1164	34.2	640	4 AAU50090	Aau50090 Propionib
8	1164	34.2	640	6 ABM46609	Abm46609 Propionib
9	823	24.1	463	6 ABM64986	Abm64986 Propionib
10	560.5	16.4	622	4 AAU60953	Aau60953 Propionib
11	560.5	16.4	622	6 ABM57472	Abm57472 Propionib
12	550	16.1	1106	6 ABU39607	Abu39607 Protein e
13	541.5	15.9	1100	4 AAU36268	Aau36268 Pseudomon
14	541.5	15.9	1100	6 ABU38432	Abu38432 Protein e
15	541	15.9	1154	6 ABU21832	Abu21832 Protein e
16	538.5	15.8	559	6 ABU23527	Abu23527 Protein e
17	529.5	15.5	598	4 AAG92273	Aag92273 C. glutami
18	529.5	15.5	617	5 ABG80322	Abg80322 C. glutam
19	521.5	15.3	1108	6 ABU41546	Abu41546 Protein e
20	509.5	15.0	601	6 ABU34768	Abu34768 Protein e
21	508	14.9	1131	6 ABU22701	Abu22701 Protein e
22	504	14.8	568	2 AAR8869	Aar8869 Maltose c
23	502.5	14.7	546	6 ABU37054	Abu37054 Protein e
24	500	14.7	590	6 ABU33845	Abu33845 Protein e
25	488	14.3	963	2 AAR84082	Aar84082 Thermosta

26	404.5	11.9	575	7 ADC96106	Adc96106 E. faeciu
27	400.5	11.8	198	4 AAU66385	Aau66385 Propionib
28	400.5	11.8	198	6 ABM62904	Abm62904 Propionib
29	373	10.9	560	5 ABM48657	Abm48657 Listeria
30	364	10.7	334	4 AAB79645	Aab79645 Corynebac
31	346	10.2	250	6 ABM64998	Abm64998 Propionib
32	342	10.0	537	6 ABU29417	Abu29417 Protein e
33	335.5	9.8	584	4 ABM69954	Abm69954 Drosophi
34	333.5	9.8	542	5 ABP25657	Abp25657 Streptoco
35	333.5	9.8	548	5 ABM49321	Abm49321 Listeria
36	332.5	9.8	549	6 ABM71064	Abm71064 Staphyloc
37	331	9.7	539	6 ABU03115	Abu03115 Alpha amy
38	328.5	9.6	549	4 AAU36999	Aau36999 Staphyloc
39	328.5	9.6	567	4 ABB63983	Abb63983 Drosophi
40	326	9.6	545	6 ADB07218	Adb07218 Alloiococ
41	323.5	9.5	546	4 AAU34014	Aau34014 Staphyloc
42	321.5	9.4	549	6 ABU15991	Abu15991 Protein e
43	319.5	9.4	541	5 ABP25656	Abp25656 Streptoco
44	317	9.3	551	4 AAG82287	Aag82287 S. epider
45	317	9.3	551	6 ABJ19033	Abj19033 Pathogen

ALIGNMENTS

RESULT 1  
AAU81945  
ID AAY81945 standard; protein; 636 AA.  
XX  
AC AAY81945;  
XX  
DT 30-JUN-2000 (first entry)  
XX  
DE Neisseria polysaccharea amylosucrase.  
XX  
KW Amylosucrase; linear alpha-1,4 glucan production; fructose production;  
KW extracellular enzyme; cyclodextrin production; pure fructose syrup.  
XX  
OS Neisseria polysaccharea.  
XX  
PN WO200014249-A1.  
XX  
PD 16-MAR-2000.  
XX  
PF 02-SEP-1998; 98WO-EP005573.  
XX  
PR 02-SEP-1998; 98WO-EP005573.  
XX  
(PLAN-) PLANTTEC BIOTECHNOLOGIE GMBH.  
XX  
PI Quanz M, Provart N;  
XX  
DR WPI; 2000-256995/22.  
XX  
DR N-PSDB; AAA07380.  
XX  
PT Novel amylosucrase polynucleotides and polypeptides obtained from  
PT Neisseria polysaccharea, used for production of -a-1,4 glucans, fructose  
PT and cyclodextrins.  
XX  
PS Claim 1; Page 25-30; 36pp; English.  
XX  
CC This sequence represents the Neisseria polysaccharea amylosucrase of the  
CC invention. The amylosucrase sequences are used for the in vitro  
CC production of linear alpha-1,4 glucans and fructose, as the enzyme is an  
CC extracellular enzyme. Amylosucrases may also be useful for the production  
CC of cyclodextrins. The amylosucrase sequences of the invention allow for  
CC the in vitro production of alpha-1,4 glucans and pure fructose syrup at  
CC low cost. Activated glucose derivatives or cofactors are not required.  
CC Immobilised cells can be at much higher densities than cells in liquid  
CC cultures, resulting in higher productivity  
XX  
SQ Sequence 636 AA;

Query Match		100.0%;	Score 3408;	DB 3;	Length 636;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 636;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1 MLPTQOVGLILQYLKTRILDIYTPQAGIEKSEDMRQFSRRMDTHFPKLMNELDSVYG 60				
Db	1 MLPTQOVGLILQYLKTRILDIYTPQAGIEKSEDMRQFSRRMDTHFPKLMNELDSVYG 60				
QY	61 NNEALLPMLEMLLAQAWQSYQNSSLKDIDIARENNPDWILSNKQVGVCYVDLFAAGDL 120				
Db	61 NNEALLPMLEMLLAQAWQSYQNSSLKDIDIARENNPDWILSNKQVGVCYVDLFAAGDL 120				
QY	121 KGLKDKIPYFQELGLTYLHLMPLFKCPBGKSDGYAVSSYRDVNPALGTIGDLREVIAL 180				
Db	121 KGLKDKIPYFQELGLTYLHLMPLFKCPBGKSDGYAVSSYRDVNPALGTIGDLREVIAL 180				
QY	181 HEAGISAVVDIFFNHTSNEHEWAQRCAAGDPLFDNFYIFPDRMPDQYDRTLREIFPDQ 240				
Db	181 HEAGISAVVDIFFNHTSNEHEWAQRCAAGDPLFDNFYIFPDRMPDQYDRTLREIFPDQ 240				
QY	241 HPGGFSQLEDGRWVWTTENSFQWDLNTSNPWFVRAMAGEMFLANLGVDIRMDAVAFIW 300				
Db	241 HPGGFSQLEDGRWVWTTENSFQWDLNTSNPWFVRAMAGEMFLANLGVDIRMDAVAFIW 300				
QY	301 KOMGTSCENLPQAHALIRAFNAVMRIAAPAVFEKSEALVHPDQVQYIGQDECQIGYNPL 360				
Db	301 KOMGTSCENLPQAHALIRAFNAVMRIAAPAVFEKSEALVHPDQVQYIGQDECQIGYNPL 360				
QY	361 QMALIMNTLATREVNLLHQALTYRHNLPEHTAMVNVYRSHDDIGWTFADEDAAYLGISGY 420				
Db	361 QMALIMNTLATREVNLLHQALTYRHNLPEHTAMVNVYRSHDDIGWTFADEDAAYLGISGY 420				
QY	421 DHRQFLNRFVNRFDGSPARGVPFQYNPSTGDCRVSGTAALVGLAQDDPHAVDRIKLTY 480				
Db	421 DHRQFLNRFVNRFDGSPARGVPFQYNPSTGDCRVSGTAALVGLAQDDPHAVDRIKLTY 480				
QY	481 SIALSTGGULPIYLIGDEVGTINDDWSQDSNKSDDSRWAHRPRYNEALYYAQRNDPSTAG 540				
Db	481 SIALSTGGULPIYLIGDEVGTINDDWSQDSNKSDDSRWAHRPRYNEALYYAQRNDPSTAG 540				
QY	541 QIYOGIRHMIAVRQSNPRFDGGRVLTENTNKKHIIIGYIRNNALLAFGNFSEYPQTVTAHT 600				
Db	541 QIYOGIRHMIAVRQSNPRFDGGRVLTENTNKKHIIIGYIRNNALLAFGNFSEYPQTVTAHT 600				
QY	601 LQAMPFKAHDLIGKTVSLNQDLTLQPYQVMMLEIA 636				
Db	601 LQAMPFKAHDLIGKTVSLNQDLTLQPYQVMMLEIA 636				
RESULT 2					
AAW90979 standard; protein; 636 AA.					
ID	AAW90979				
AC	AAW90979;				
XX	21-JUL-2000 (first entry)				
DT	N. polysacchara amylosucrase protein.				
XX	Alpha-1,6-branched alpha-1,4-glucan; branching enzyme; binder; carrier;				
KW	transgenic plant; flavor; perfume; packaging material; papermaking;				
KW	ultra-violet light adsorber; starch; textile; wetting agent;				
KW	amylsucrase.				
OS	Neisseria polysacchara.				
XX	WO200022140-A1.				
PD	20-APR-2000.				
PF	08-OCT-1999; 99WO-EP007562.				
PR	09-OCT-1998; 98DE-01046635.				

PR	27-MAY-1999; 99DE-01024342.				
XX	(PLAN-) PLANTTEC BIOTECHNOLOGIE GMBH.				
PA	(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.				
PA	Buettcher V, Quanz M;				
PI	WPI; 2000-317992/27.				
DR	N-PSDB; AAA11732.				
XX	New nucleic acid encoding a branching enzyme, useful for in vitro				
PT	synthesis of branched glucans and to prepare transgenic plants producing				
PT	modified starch.				
XX	Disclosure; Page 102-104; 115pp; German.				
PS	This invention describes a novel nucleic acid (I) isolated from Neisseria				
CC	which encodes a branching enzyme (II). (I) is used for recombinant				
CC	production of (II) subsequently used in the in vitro production of alpha-				
CC	1,6-branched alpha-1,4-glucans. It is also used to prepare transgenic				
CC	plants that produce starches with modified properties. (III) are used as				
CC	binders for tablets, carriers for pharmaceuticals, flavors and perfumes				
CC	and powdered additives, packaging materials, ultra-violet light adsorbers				
CC	in sunscreens and also for any of the usual applications of starch in				
CC	foods, papermaking, as textile size, in soil stabilization, as wetting				
CC	agent for agricultural chemicals, as polymer additives etc. Fragments of				
CC	(I) are useful as PCR primers and antisense molecules or ribozymes for				
CC	inhibiting expression of (I), and the regulatory region of (II) can be				
CC	used to control expression of heterologous sequences in host cells. (I)				
CC	provides an inexpensive method for producing alpha-1,6-branched alpha-1,4				
CC	-glucans (Iii), producing products that can be tailored for particular				
CC	applications, particularly by controlling the degree of branching. Starch				
CC	from transgenic plants has increased gel strength; reduced phosphate				
CC	content; reduced peak viscosity; lower pasting temperature and granule				
CC	size and/or altered sidechain distribution. This sequence represents an				
CC	amylsucrase isolated from Neisseria polysacchara which is described in				
CC	the method of the invention				
XX	Sequence 636 AA;				
SQ					
Query Match					
Best Local Similarity 100.0%; Score 3408; DB 3; Length 636;					
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1 MLPTQOVGLILQYLKTRILDIYTPQAGIEKSEDMRQFSRRMDTHFPKLMNELDSVYG 60				
Db	1 MLPTQOVGLILQYLKTRILDIYTPQAGIEKSEDMRQFSRRMDTHFPKLMNELDSVYG 60				
QY	61 NNEALLPMLEMLLAQAWQSYQNSSLKDIDIARENNPDWILSNKQVGVCYVDLFAAGDL 120				
Db	61 NNEALLPMLEMLLAQAWQSYQNSSLKDIDIARENNPDWILSNKQVGVCYVDLFAAGDL 120				
QY	121 KGLKDKIPYFQELGLTYLHLMPLFKCPBGKSDGYAVSSYRDVNPALGTIGDLREVIAL 180				
Db	121 KGLKDKIPYFQELGLTYLHLMPLFKCPBGKSDGYAVSSYRDVNPALGTIGDLREVIAL 180				
QY	181 HEAGISAVVDIFFNHTSNEHEWAQRCAAGDPLFDNFYIFPDRMPDQYDRTLREIFPDQ 240				
Db	181 HEAGISAVVDIFFNHTSNEHEWAQRCAAGDPLFDNFYIFPDRMPDQYDRTLREIFPDQ 240				
QY	241 HPGGFSQLEDGRWVWTTENSFQWDLNTSNPWFVRAMAGEMFLANLGVDIRMDAVAFIW 300				
Db	241 HPGGFSQLEDGRWVWTTENSFQWDLNTSNPWFVRAMAGEMFLANLGVDIRMDAVAFIW 300				
QY	301 KOMGTSCENLPQAHALIRAFNAVMRIAAPAVFEKSEALVHPDQVQYIGQDECQIGYNPL 360				
Db	301 KOMGTSCENLPQAHALIRAFNAVMRIAAPAVFEKSEALVHPDQVQYIGQDECQIGYNPL 360				
QY	361 QMALIMNTLATREVNLLHQALTYRHNLPEHTAMVNVYRSHDDIGWTFADEDAAYLGISGY 420				
Db	361 QMALIMNTLATREVNLLHQALTYRHNLPEHTAMVNVYRSHDDIGWTFADEDAAYLGISGY 420				
QY	421 DHRQFLNRFVNRFDGSPARGVPFQYNPSTGDCRVSGTAALVGLAQDDPHAVDRIKLTY 480				

```
Db 421 DHRQFLNRFVNRFDGSGFARGVPEQYNPSTGDCRVSGTAALVGLAQDDPHAVDRIKLLY 480
QY 481 SIALSTGGLPLIYLGDVGTLLNDDWSQDSNKSDDSRWAHRPRYNEALYAQRNDPSTAG 540
Db 481 SIALSTGGLPLIYLGDVGTLLNDDWSQDSNKSDDSRWAHRPRYNEALYAQRNDPSTAG 540
QY 541 QIYQGLRHMIAVRQSNRPFDDGRLVTENTNNKHIIIGYTRNNALLAFGNFSEYPQVTYAH 600
Db 541 QIYQGLRHMIAVRQSNRPFDDGRLVTENTNNKHIIIGYTRNNALLAFGNFSEYPQVTYAH 600
QY 601 LQAMPFKAHDLIGKTVSLNODLTLPYQVMMLEIA 636
Db 601 LQAMPFKAHDLIGKTVSLNODLTLPYQVMMLEIA 636
```

RESULT 3  
AAY87831  
ID AAY87831 standard; protein; 636 AA.

```
XX AC AAY87831;
XX DT 01-SEP-2000 (first entry)
XX DE N. polysaccharea amylsucrase protein.
XX KM Polysglucan; amylsucrase; active ingredient carrier; cosmetic;
XX OS agrochemical; food; food additive; flavor; perfume.
XX OS Neisseria polysaccharea.
XX PN WO200022155-A2.
XX PD 20-APR-2000.
XX PF 07-OCT-1999; 99WO-EP007518.
XX PR 09-OCT-1998; 98DE-01046492.
XX PA (AVET ) AVENTIS RES & TECHNOLOGIES GMBH & CO KG.
XX PI Gallert K, Bengs H, Simandl C;
XX DR WPI; 2000-318003/27.
```

XX PT New polysglucan and derivatives, useful e.g. as carriers for  
PT pharmaceuticals, agrochemicals and cosmetics, prepared using polysglucan  
PT sucrose and an additional biogenic compound.

XX PS Claim 3; Page 24-26; 26pp; German.

XX CC This invention describes a novel polysglucan and its derivatives (A) which  
CC is produced using a polysglucan sucrose or amylsucrase in the presence of  
CC at least one other biogenic material (I). (A) are used; (i) as carriers  
CC for active ingredients especially for pharmaceutical, cosmetic and  
CC agrochemical compounds; (ii) as food or food additives; and (iii) as  
CC carriers for flavors or perfumes. They may also be used as starting  
CC materials for further chemical modification. The activity of amylsucrase  
CC is not adversely affected by presence of (I), particularly other enzymes,  
CC and the combination of enzymes allows a very wide range of modified  
CC products to be prepared. In vitro reaction facilitates production of  
CC reproducible products of consistent quality and value, with properties  
CC (e.g. molecular weight, primary structure, polydispersity etc.) tailored  
CC for specific applications. This sequence represents a Neisseria  
CC polysaccharea amylsucrase which is described in the method of the  
CC invention

XX SQ Sequence 636 AA;

Query Match 100.0%; Score 3408; DB 3; Length 636;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MLTPTOQVGLIQLYLRILDIYTPQAGIEKSEDMQFSRMDTHFKLMLNELDSVYG 60
Db 1 MLTPTOQVGLIQLYLRILDIYTPQAGIEKSEDMQFSRMDTHFKLMLNELDSVYG 60
QY 61 NNEALLPMLMLLAQAMQSYSQRNSSLKDDIDARENNPDWILSNKQGVGVCLDLFAGDL 120
Db 61 NNEALLPMLMLLAQAMQSYSQRNSSLKDDIDARENNPDWILSNKQGVGVCLDLFAGDL 120
QY 121 KGLKDKIPYFOELGLTYLHMLPLEKCEKSGDGYAVSSYRDVNPALGTIGDLREVIAL 180
Db 121 KGLKDKIPYFOELGLTYLHMLPLEKCEKSGDGYAVSSYRDVNPALGTIGDLREVIAL 180
QY 181 HEAGISAVVDPIFNHTSNEHEWAQRCAGDPLFDNFYYIFPDRMPDQYDRTLREIFPDQ 240
Db 181 HEAGISAVVDPIFNHTSNEHEWAQRCAGDPLFDNFYYIFPDRMPDQYDRTLREIFPDQ 240
QY 241 HPGGFSQLEDGRWWTTFNSFQWDLNTSNPWVFRAMGEMLFLANLGVDTLRMDAVAFIW 300
Db 241 HPGGFSQLEDGRWWTTFNSFQWDLNTSNPWVFRAMGEMLFLANLGVDTLRMDAVAFIW 300
QY 301 KOMGTSCENLPQAHALIRAFNAVMRIAPAVFPKSEALVHPDQVQYIGQDECQIGYNPL 360
Db 301 KOMGTSCENLPQAHALIRAFNAVMRIAPAVFPKSEALVHPDQVQYIGQDECQIGYNPL 360
QY 361 QMALLMNTLATREVNLIHQALTYRHNLPEHTAWNVYVRSHDDIGWTFADEDAAYLGISGY 420
Db 361 QMALLMNTLATREVNLIHQALTYRHNLPEHTAWNVYVRSHDDIGWTFADEDAAYLGISGY 420
QY 421 DHRQFLNRFVNRFDGSGFARGVPEQYNPSTGDCRVSGTAALVGLAQDDPHAVDRIKLLY 480
Db 421 DHRQFLNRFVNRFDGSGFARGVPEQYNPSTGDCRVSGTAALVGLAQDDPHAVDRIKLLY 480
QY 481 SIALSTGGLPLIYLGDVGTLLNDDWSQDSNKSDDSRWAHRPRYNEALYAQRNDPSTAG 540
Db 481 SIALSTGGLPLIYLGDVGTLLNDDWSQDSNKSDDSRWAHRPRYNEALYAQRNDPSTAG 540
QY 541 QIYQGLRHMIAVRQSNRPFDDGRLVTENTNNKHIIIGYTRNNALLAFGNFSEYPQVTYAH 600
Db 541 QIYQGLRHMIAVRQSNRPFDDGRLVTENTNNKHIIIGYTRNNALLAFGNFSEYPQVTYAH 600
QY 601 LQAMPFKAHDLIGKTVSLNODLTLPYQVMMLEIA 636
Db 601 LQAMPFKAHDLIGKTVSLNODLTLPYQVMMLEIA 636
```

RESULT 4  
AAB70883  
ID AAB70883 standard; protein; 856 AA.

XX AC AAB70883;

XX DT 12-JUL-2001 (first entry)

XX DE Expression vector pGEX-4T-1-AmsU5 containing amylsucrase.

XX KM Amylosucrase; EC 2.4.1.4; fusion protein; GST; glutathione-S-transferase;  
XX poly(1,4-alpha-glucan); film production; food additive; cyclodextrin.

XX OS Unidentified.

XX PN WO200125449-A2.

XX PD 12-APR-2001.

XX PF 04-OCT-2000; 2000WO-EP009695.

XX PR 07-OCT-1999; 99DE-01048408.

XX PA (AXIV-) AXIVA GMBH.

XX PI Bengs H, Polakowski T, Held A, Gallert K;

XX DR WPI; 2001-328330/34.



DR N-PSDB; AAF61712.

XX Amylosucrase immobilized as fusion protein with anchoring sequence,  
PT useful in continuous preparation of poly(1,4-alpha-glucan) from sucrose.  
XX  
PS Claim 2; Page 34-38; 38pp; German.

CC This invention describes a novel amylosucrase (AS), immobilizable on a  
CC solid phase, which comprises a fusion protein (FP) of functional units of  
CC AS, an anchoring sequence, and optionally additional auxiliary sequences.  
CC The invention also describes (1) nucleic acid (I) encoding FP; (2)  
CC expression vector containing (I) and able to express FP in a host cell;  
CC Escherichia coli containing the vector of (2); (3) anchoring sequence, or  
CC its functional variants or fragments, of at least 8 nucleotides that  
CC encodes an epitope, a high-affinity binding partner or GST (glutathione-S-  
CC -transferase); (4) solid phase for immobilizing AS comprising glutathione  
CC -Sepharose; (5) combined, stable catalyst (A) comprising FP immobilized  
CC on Sepharose for production of poly(1,4-alpha-glucan) (II); and (6)  
CC biocatalytic production of (II). AS is used in production of poly(1,4-  
CC alpha-glucans) (II), useful for producing films, as food additives, as  
CC starting materials for cyclodextrins and as auxiliaries in pharmaceutical  
CC formulations. Immobilized AS makes possible efficient, inexpensive and  
CC continuous production of poly(1,4-alpha-glucans) (II), and it can be used  
CC repeatedly. Compared with known methods, specificity is improved  
CC (increased yield of (II) and reduced formation of palatinose) and  
CC reaction is complete within 24 hours, compared to 48-72 hours for batch  
CC methods. This sequence represents the expression vector construct pGEX-4T  
CC -1-AmsUs which contains the amylosucrase (EC 2.4.1.4) described in the  
CC invention

XX  
XX Sequence 856 AA;

Query Match 98.6%; Score 3360; DB 4; Length 856;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 626; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 LILQYKTRIDIDYTPQKAGIEKSEDMQFQSRMDTHFPKLMNELDSVYGNNEALLPML 69  
:|||||  
DB 230 MILQYKTRIDIDYTPQKAGIEKSEDMQFQSRMDTHFPKLMNELDSVYGNNEALLPML 289

QY 70 EMLLAQAWQSQGRNSLSKDIDIARENNPDWILSNKQVGVYVDLFAAGDLKGLKDKIPY 129  
|||||  
DB 290 EMLLAQAWQSQGRNSLSKDIDIARENNPDWILSNKQVGVYVDLFAAGDLKGLKDKIPY 349

QY 130 FOELGLTYLHMLPFCPEGKSDGYAVSSYRDVNPALGTIGDLREVIALHEAGISAVV 189  
|||||  
DB 350 FOELGLTYLHMLPFCPEGKSDGYAVSSYRDVNPALGTIGDLREVIALHEAGISAVV 409

QY 190 DFIFNHTSNEHEWAQRCAGDPLFDNFYIIFPDRRMPDQYDRTLREIFPDQHPGFSQLE 249  
|||||  
DB 410 DFIFNHTSNEHEWAQRCAGDPLFDNFYIIFPDRRMPDQYDRTLREIFPDQHPGFSQLE 469

QY 250 DGRWVWTTENSFOQMDLINSNPVFRAMAGEMFLANLGVLDILRMDAVAFIWKMGTSCE 309  
|||||  
DB 470 DGRWVWTTENSFOQMDLINSNPVFRAMAGEMFLANLGVLDILRMDAVAFIWKMGTSCE 529

QY 310 LPOAHALIRAFNAVRIAAPAVFCKSEAIYHPDQVQYIGQDECQIGINPLQMALMNTL 369  
|||||  
DB 530 LPOAHALIRAFNAVRIAAPAVFCKSEAIYHPDQVQYIGQDECQIGINPLQMALMNTL 589

QY 370 ATREVNLLHQALTYRHNLPEHTAMVNVYRSHDDIGWTFADEDAAYLGISGYDHRQFLNRF 429  
|||||  
DB 590 ATREVNLLHQALTYRHNLPEHTAMVNVYRSHDDIGWTFADEDAAYLGISGYDHRQFLNRF 649

QY 430 FVNRFDGSFARGVPEQYNPSTGDCRVSGTAALVGLAQDDPHAVDRIKLLYSIALSTGGL 489  
|||||  
DB 650 FVNRFDGSFARGVPEQYNPSTGDCRVSGTAALVGLAQDDPHAVDRIKLLYSIALSTGGL 709

QY 490 PLIYLGDEVGTINDDDWSQDSNKSDDSRWAHPRRYNEALYAQRNDPSTAGQIYQGLRHM 549  
|||||  
DB 710 PLIYLGDEVGTINDDDWSQDSNKSDDSRWAHPRRYNEALYAQRNDPSTAGQIYQGLRHM 769

QY 550 IAVRQSNPRFDGRLVTENTNKHIIIGYIRNNALLAFGNFSEYPGYVTAHTLQAMPFKAH 609

DB 770 IAVRQSNPRFDGRLVTENTNKHIIIGYIRNNALLAFGNFSEYPGYVTAHTLQAMPFKAH 829  
|||||

QY 610 DLIGKTVSLNQDLTLQPYQVWMLAIA 636  
|||||

DB 830 DLIGKTVSLNQDLTLQPYQVWMLAIA 856  
|||||

RESULT 5  
AAR88386  
ID AAR88386 standard; protein; 614 AA.  
XX  
AC AAR88386;  
XX  
DT 15-OCT-1996 (first entry)  
XX  
DE Neisseria polysaccharea amylosucrase.  
XX  
KW Amylosucrase; enzyme; amylose, linear 1,4-glucan; transgenic plant;  
EC-2.4.1.4; crop improvement.  
XX  
OS Neisseria polysaccharea.  
XX  
PN W09531553-A1.  
XX  
PD 23-NOV-1995.  
XX  
PF 18-MAY-1995; 95WO-EP001893.  
XX  
PR 18-MAY-1994; 94DE-04417879.  
PR 22-DEC-1994; 94DE-04447388.  
XX  
PA (GENB-) INST GENBIOLOGISCHE FORSCHUNG.  
XX  
PI Kossmann J, Buettcher V, Welsh T;  
XX  
DR WPI; 1996-010938/01.  
DR N-PSDB; AAT08960.  
XX  
PT DNA encoding amylo: sucrose from Neisseria polysaccharea - for prodn. of  
PT linear 1,4-glucan(s), esp. amylose, from sucrose.  
XX  
PS Claim 1; Page 39-42; 56pp; English.  
XX  
CC This sequence encodes an amylosucrase which allows the synthesis of  
CC linear alpha-1,4-glucans from the substrate sucrose by bacteria, fungi  
CC and plants, or in cell-free systems. This sequence may be expressed  
CC recombinantly. Typical applications of amylose are production of films  
CC and fibres for use in food, textile and glass fibre industries or paper  
CC manufacture, etc. Amylosucrase may be used in the low cost production of  
CC pure fructose syrups and cyclodextrin

XX  
XX Sequence 614 AA;

Query Match 74.5%; Score 2540; DB 2; Length 614;  
Best Local Similarity 82.1%; Pred. No. 5.1e-239;  
Matches 494; Conservative 13; Mismatches 43; Indels 52; Gaps 6;

QY 1 MLTPYQVGLILQYKTRIDIDYTPQKAGIEKSEDMQFQSRMDTHFPKLMNELDSVYG 60  
|||||  
DB 1 MLTPYQVGLILQYKTRIDIDYTPQKAGIEKSEDMQFQSRMDTHFPKLMNELDSVYG 60

QY 61 NNEALLPMLMELLAQAWQSQGRNSLSKDIDIARENNPDWILSNKQVGVYVDLFAAGDL 120  
|||||  
DB 61 NNEALLPMLMELLAQAWQSQGRNSLSKDIDIARENNPDWILSNKQVGVYVDLFAAGDL 120

QY 121 KGLKDKIPYFQELGLTYLHMLPFCPEGKSDGYAVSSYRDVNPALGTIGDLREVIAL 180  
|||||  
DB 121 KGLKDKIPYFQELGLTYLHMLPFCPEGKSDGYAVSSYRDVNPALGTIGDLREVIAL 180

QY 181 HEAGISAVYDFIFNHTSNEHEWAQRCAGDPLFDNFYIIFPDRRMPDQYDRTLREIFPDQ 240  
|||||  
DB 181 HESHFRRRRFFYLQPHLQRT-RMAQRC-AGDPLFDNFYIIFPDRRMPDQYDRTLREIFPDQ 238

QY	241	HFGGFSQLEDGRWVMTTNSFQMDLNTSNPWF---RAMAGEMFLANLGVDILRMDAVA	297
		:           :         :	
Dd	239	HFGGFSQLEJEDGRWVMTTFNSFQMDLNTSNPWFAQWRKACCSLPWTALTSVCVMRLPLFG	298
QY	298	FIMKOMGTSCENLPOAHALIRAFNAVMRIAPAVEFKSEAIIVHPDQVVQYIGQDECOIGY	357
		:	
Dd	299	NKMWGQAKTCT---AAHALIRAFNAVMRIAPAVEFKSEAIIVHPDQVVQYIGQDECOIGY	354
QY	358	NPLQMALLMNTLATREVNLLHQALTYSRHNLPEHTAVWNVYVRSHDDIGWTEADEDAAYLGI	417
Dd	355	NPLQMALLMNTLATREBNLLHQUALTYRNHLPEHTAVWNVYVRSHDDIGWTEADEDAAYLGI	414
QY	418	SGVDHRQFLNREFVNRFDGSFARGVPFOYNPSTGDCRVSGTAALVTGLAQDBHAVDRIK	477
Dd	415	SGVDHRQFLNREFVNRFDGTFARGVPFQYNPSTGDCRVSGTAALVTGLAQDBHAVDRIK	474
QY	478	LITYIALSTGGCLPLYLYLGDEVGTLNDDEWSQDSNKSDDSRWahrPRyNEALYAQRNDPST	537
Dd	475	LITYIAIISTGGCLPLYLYLGDEVGTLNDDEWCQAIRA-----TT	512
QY	538	AAG-----QIYGGLRMIIA VRQS N PR FD G RL VT F NT NN K H I IG	576
Dd	513	AYGPSPSALO R SP V RAT ERS V DR SR Q I Y O GL RH M IA VR QS NP RF DG RL VT F NT NN K H I IG	572
QY	577	YT 578	
Dd	573	YT 574	

RESULT 6  
AAR88633  
ID AAR88633 standard; protein; 614 AA.  
XX  
XX AAR88633;  
AC  
XX 20-JUN-1996 (first entry)  
DT  
XX  
DE Neisseria polysaccharea amylsucrase.  
XX  
XX Amylosucrase; bacteria; fungi; plants; detection; transformation; linear;  
XX alpha-1,4-glucans; amylose; sucrose; colourless; odourless; tasteless;  
XX non-toxic; biodegradable; self-sustaining; films; fibres; textiles;  
XX paper-making; glass-fibre; tablet binder; food thickener; sound proofing;  
XX flow properties; paraffin oils; organic compound inclusion;  
XX chromatographic separation; cyclodextrins.  
XX  
XX Neisseria polysaccharea.  
OS  
XX  
XX DE4417879-A1.  
PN  
XX  
XX 23-NOV-1995.  
PD  
XX  
XX 18-MAY-1994; 94DE-04417879.  
PF  
XX  
XX 18-MAY-1994; 94DE-04417879.  
PR  
XX  
XX (GENB-) INST GENBIOLOGISCHE FORSCHUNG.  
PA  
XX  
XX Kossmann J, Buettcher V, Welsh T;  
PI  
XX  
XX Kossmann J, Buettcher V, Welsh T;  
PI  
XX  
XX MPI; 1996-000447/01.  
DR  
XX  
XX N-PSDB; AAT11179.  
DR  
XX  
XX New DNA sequence encoding amyl:sucrase of Neisseria - and transformed  
PT plant, bacteria and fungi able to produce linear alpha-1,4-glucan(s),  
PT esp. amylose, in practically pure form.  
PT  
XX  
XX  
PS  
XX  
XX Claim 5; Page 23-27; 42pp; German.  
XX  
XX The N. polysaccharea DNA sequence AAT11179, which encodes AAR88633  
CC amylsucrase (ASA), can be used to produce bacteria, fungi and plants  
CC that express ASA, and to detect and isolate related DNA from other

CC organisms. Transformed plants which express ASA are able to produce  
CC linear alpha-1,4-glucans, specifically amylose from sucrose, which can be  
CC used to produce colourless, odourless, tasteless, nontoxic,  
CC biodegradable, self-sustaining films or fibres, e.g. for use in the food,  
CC textile, paper- making and glass-fibre industries. Amylose can also be  
CC used as a binder for tablets, thickener for food, in sound proofing  
CC panels, to improve flow properties in paraffin-based oils, for inclusion  
CC of organic cpds., in chromatographic sepn. and as a starting material for  
CC cyclodextrins  
XX  
SQ Sequence 614 AA;

Query Match	74.5%;	Score 2540;	DB 2;	Length 614;
Best Local Similarity	82.1%;	Pred. No. 5.1e-239;		
Matches 494;	Conservative 13;	Mismatches 43;	Indels 52;	Gaps 6;

QY	1	MLTPTQOVGLILOYLKTRILDIYTP	EOBAGIEKSEDMRQFSRRMDTHFP	PKLMNELDSVYG	60
Db	1	MLTPTQOVGLILOYLKTRILDIYTP	EOBAGIEKSEDMRQFSRRMDTHFP	PKLMNELDSVYG	60
QY	61	NNEALLPMLLEMLLAQAMQS	SOQNSSLKDIDIARENNPDWILSNKQ	VGVCYVDL	120
Db	61	NNEALLPMLLEMLLAQAMQS	SOQNSSLKDIDIARENNPDWILSNKQ	VGVCYVDL	120
QY	121	KGKDKPIFYFOELGLTYLHMLP	FKCPEKSGDGYAVSSYRDVNPALGTI	GDLEVI	180
Db	121	KGKDKPIFYFOELGLTYLHMLP	FKCPEKSGDGYAVSSYRDVNPALGTI	GDLEVI	180
QY	181	HEAGISAVVDPIFNHTSNEHEWA	QRCAGADPLFDNFYIIFPDRMPQYDRT	LEIFPDQ	240
Db	181	HEAGISAVVDPIFNHTSNEHEWA	QRCAGADPLFDNFYIIFPDRMPQYDRT	LEIFPDQ	240
QY	241	HGGFSQLEDGRWWTTFNSFQMD	LNYSNPWF---RAMAGEMLFLANL	GDILRMDAVA	297
Db	241	HGGFSQLEDGRWWTTFNSFQMD	LNYSNPWF---RAMAGEMLFLANL	GDILRMDAVA	297
QY	298	FIWKOMGTSCEMLPQAHALIRAF	NAVMRIAAPAVFPKSEATVHPDQV	YIGQDECQIGY	357
Db	298	FIWKOMGTSCEMLPQAHALIRAF	NAVMRIAAPAVFPKSEATVHPDQV	YIGQDECQIGY	357
QY	358	NPLQOMALLMNTLATREVNLI	HQALTYRHNLPEHTAMVNVYVRSHDDI	IGWTEAD	417
Db	358	NPLQOMALLMNTLATREVNLI	HQALTYRHNLPEHTAMVNVYVRSHDDI	IGWTEAD	417
QY	418	SGYDHRQFLNRFVNRFDGSP	ARGVPEQYNPSTGDCRVSGTAAL	VGLAQDDPHAVDRIK	477
Db	418	SGYDHRQFLNRFVNRFDGSP	ARGVPEQYNPSTGDCRVSGTAAL	VGLAQDDPHAVDRIK	477
QY	478	LLYSIALSTGGLPLIYL	LGDEVTLDDDDWSQDSNKS	DDSRWAHRPRYNEALYAQRND	537
Db	478	LLYSIALSTGGLPLIYL	LGDEVTLDDDDWSQDSNKS	DDSRWAHRPRYNEALYAQRND	537
QY	538	AAQ-----QIYQGLRH	MIAYRQSNPRFDGRLVT	FNTNNKHIIIG	576
Db	538	AAQ-----QIYQGLRH	MIAYRQSNPRFDGRLVT	FNTNNKHIIIG	576
QY	577	YI 578			
Db	577	YI 578			
QY	573	YI 574			
Db	573	YI 574			

RESULT 7	
AAU50090	
ID	AAU50090 standard; protein; 640 AA.
XX	
AC	AAU50090;
XX	
DT	27-FEB-2002 (first entry)
XX	
DE	Propionibacterium acnes immunogenic protein #10986.
XX	
KW	SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW	uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;









QY	290	ILRMDAVAFIWKQMTSCENLPQAHALIRAFNAVMRIAAPAVFEKSEATVHPDQVQYIG	349
Db	95	VERLDIAIAFIWKLGITNCQNLPEIHDTQSLRQAIRIVAPAVAFMADAIYGPDDLTYGYG	154
QY	350	QDE-----CQIGYNPLQMALMTLTATREYNLLHQALTYRHNLPHTAMVNYVRSHDDIG	404
Db	155	RGRHWGKVCMDMIYHNSIMVQLMSALATFDVSLMETALSRTPDKPSTTTWATYARCHDDIG	214
QY	405	WTFADEDAAYLGISGYDHRQFLNRFVNRFDGSFARGVPFQYNPSTGDCRVSGTAALVG	464
Db	215	WTVDDADARKTGCLDPVAHRQPLSDFYSGTFPGSFARGLVFQDNPTVTGDRIRISGSLASLAG	274
QY	465	----LAQDDPFAVD---RIKLYSIALSTGGPLIYLGDVEVGTLLNDDNSQDSNKSDDS	516
Db	275	LESALIESDDPAGVDAAIARIIVMLHTAILGYGCVPLIMGDEVGMILN-DWQRPDEGHADDN	333
QY	517	RNAHRPRYNEALYAQRN-DPSTAAGQIYQGLRHMIAVRQSNPRFDGG-RLVTENTNNKHI	574
Db	334	RWVHRPMMNWSMVKQAHAEPHSVPGRIWNGVRRAINARHRSPEFHAASVDTVVLPSPHRKV	393
QY	575	IGYIR---NNALLAFGNFSEYPPQVTVAHTLQAMPFKAHDLIQKTVSINODLT-----LQ	626
Db	394	IMWGRPHPEGRMIELYNISEHEWVFPMETLRS---ELDDVVTETLLRGFDYDLTPMNLRLA	450
QY	627	PYQVMWL 633	
Db	451	PYECLMWL 457	
RESULT 10			
AAU60953	ID	AAU60953 standard; protein; 622 AA.	
XX	AC	AAU60953;	
XX	DT	27-FEB-2002 (first entry)	
XX	DE	Propionibacterium acnes immunogenic protein #21849.	
XX	XX	SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;	
KW	KW	uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;	
KW	KW	inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;	
KW	KW	dermatological; osteopathic; neuroprotectant.	
XX	OS	Propionibacterium acnes.	
XX	PN	WO200181581-A2.	
XX	PD	01-NOV-2001.	
XX	XX	20-APR-2001; 2001WO-US012865.	
XX	XX	21-APR-2000; 2000US-0199047P.	
PR	PR	02-JUN-2000; 2000US-0208841P.	
PR	PR	07-JUL-2000; 2000US-0216747P.	
XX	XX	(CORI-) CORIXA CORP.	
XX	XX	Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;	
PI	PI	L'maisonneuve J, Zhang Y, Jen S, Carter D;	
XX	XX	WPI; 2001-616774/71.	
DR	DR	N-PSDB; AAS59613.	
XX	XX	Propionibacterium acnes polypeptides and nucleic acids useful for	
PT	PT	vaccinating against and diagnosing infections, especially useful for	
PT	PT	treating acne vulgaris.	
XX	XX	Example 1; SEQ ID NO 22148; 1069pp; English.	
XX	XX	Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic	
CC	CC	polypeptides. The proteins and their associated DNA sequences are used in	
CC	CC		

the treatment, prevention and diagnosis of medical conditions caused by P. acnes. The disorders include SAPHO syndrome (synovitis, acne, pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for P. acnes proteins. These antibodies can be used to downregulate expression and activity of P. acnes polypeptides and therefore treat P. acnes infections. The antibodies may also be used as diagnostic agents for determining P. acnes presence, for example, by enzyme linked immunosorbent assay (ELISA). Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIFO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences)

[illegible]

XX	DE	Propionibacterium acnes predicted ORF-encoded polypeptide #22148.
XX	KM	Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX	KW	immunostimulant; immune response; vaccine.
XX	OS	Propionibacterium acnes.
XX	PN	WO2003033515-A1.
XX	PD	24-APR-2003.
XX	PF	11-OCT-2002; 2002WO-US032727.
XX	PR	15-OCT-2001; 2001US-00978825.
XX	PA	(CORI-) CORIXA CORP.
PI	PI	Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI	PI	Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI	PI	Barth B, Valliee-Douglas J;
DR	DR	WPI; 2003-381789/36.
DR	DR	N-PSDB; ACF64542.
XX	PT	New Propionibacterium acnes polypeptides and polymucleotides encoding the
PT	PT	polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT	PT	or for stimulating an immune response specific for a P. acnes protein.
XX	PS	Example 1; SEQ ID NO 22148; 1481pp; English.
CC	CC	The invention relates to an isolated polymucleotide (ACF64435-ACF64733)
CC	CC	encoding a Propionibacterium acnes protein. The invention also relates to
CC	CC	polypeptides encoded by the polymucleotides (ABM35624-ABM64536) and to
CC	CC	immunogenic fragments of P. acnes polypeptides. The invention
CC	CC	additionally encompasses expression vectors and host cells comprising a
CC	CC	polymucleotide of the invention; antibodies against polypeptides of the
CC	CC	invention; fusion proteins comprising a polypeptide of the invention; a
CC	CC	method for stimulating an immune response specific for a P. acnes
CC	CC	polypeptide and an isolated T cell population comprising T cells prepared
CC	CC	via this method; a vaccine composition (comprising P. acnes polypeptides,
CC	CC	polymucleotides, antibodies, fusion proteins, T cell populations, or
CC	CC	antigen-presenting cells that express the polypeptide); a method and kit
CC	CC	for detecting or determining the presence or absence of P. acnes in a
CC	CC	patient; and a method for inhibiting the development of P. acnes in a
CC	CC	patient. The P. acnes polypeptides, polymucleotides, antibodies, fusion
CC	CC	proteins, T cell populations or antigen-presenting cells that express the
CC	CC	polypeptides are useful for diagnosing, preventing or treating acne
CC	CC	vulgaris, or for stimulating an immune response specific for a P. acnes
CC	CC	protein. The polymucleotides can also be used as probes or primers for
CC	CC	nucleic acid hybridisation. The vaccine composition is useful for the
CC	CC	stimulation of an immune response against P. acnes, or for treating acne,
CC	CC	and the kit is useful for performing a diagnostic assay. The present
CC	CC	sequence represents a polypeptide predicted to be encoded by an ORF (open
CC	CC	reading frame) contained within the P. acnes polymucleotides of the
CC	CC	invention. Note: The sequence data for this patent did not form part of
CC	CC	the printed specification, but was obtained in electronic format directly
CC	CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	SEQ	Sequence 622 AA;
QY	Query Match	16.4%; Score 560.5; DB 6; Length 622;
QY	Best local Similarity	28.6%; Pred. No. 3.5e-45;
Db	Matches 162; Conservative	85; Mismatches 220; Indels 99; Gaps 22
QY	118	GDLKGLKD KIPYFOELGLYLHMLPLFKCPEGSKSDGYAVSSYRDVNPALGTIGDLREVI 177
Db	104	GDFKGLTGKL DYLOWLGVDCIMLPFPYDSP--LHDGGYDIRDYRWIRELGITIEDPKVFL 161
QY	178	AALHEAGISAVVD FIFNHTSNHEHWAQRCAAG-DPLFDNFYYTFPDRRMPDQYDRTLREI 236
Db	162	DAHDRGLRVIIDFVNMTSDSHPWFOSSRADDPGYGN-YTVWSD---TDEAYSDARII 217

QY 237 FPDQHPGGFSQLEDGRWV--TTFNSFQW-----DLNYSNPWVFRAMAGEMLFLANLG 287  
 Db 218 FCD-----TEDSNMSWDSQKQFYWHRFHHQPDNLNFEEPRVMEEMLDVRFWMDLG 269  
 QY 288 VDLRMDAVAFIWKQMGTSCENLPOAHALIRAFNAWMRIAAPAVEFKSEALVHPDQVOY 347  
 Db 270 IDGFRLDVAPYLLIEAGETNGCENLPGTHKILKQLRAMVDEEYFGRILICEANQWPDVVEY 329  
 QY 348 IGQ-DECOIGYNPLQWALIMNTLATREVNLLHQALTYRRHNLPEHTAWNVYRSHDDIGWT 406  
 Db 330 FGNNGDECQMAFHFPMVPRLYMGRLRGSRRECISEILATPHIPDGCQWGTFLRNHDELTLLE 389  
 QY 407 F-ADEDAAYLGISGYDHRQFLNRFVNRFPDGSFARGVPFQYNP-STGDCRVSGTAALVG 464  
 Db 390 MYTEEDRRHYM-----EETAPESRMRCNL-GIRRLISP 421  
 QY 465 LAQDDPHAVDRIKLYSLALSTGGPLIYLGDEVGTLNDDWSQDSN-----KSDDSRW 518  
 Db 422 LVDNDD--AKIRLLNAMLLSLPGSPVLYYDDEIG-MGDDPWLPRDGVRTPMQMDDET 477  
 QY 519 A-----HPRRY-NEALYAQRNDPSTAAGQIYQGLRHMIAVRQSNPR 558  
 Db 478 AGFSTALPEDFHLPLIRTFGHDPEHVNV--RQMDDPSS---LLVWTRAMLGIRRHFPV 531  
 QY 559 FDGGRLVTFNTNNKHIIIGYIRNN---ALLAFGNFSEYPQVTVAHTLQAMPFKAHDLIGK 615  
 Db 532 FGTGEFTDLGGPDMAVMSFLRNHEHETVLCANFSDTERMVALLPQFAGMTGSSLLIHQ 591  
 QY 616 TVS-LNQDLTLQ---PYQVMWLEIA 636  
 Db 592 DAQPVKADGTLSPVLPMPYGYRMLQVS 617  
 RESULT 12  
 ABU39607  
 ID ABU39607 standard; protein; 1106 AA.  
 AC ABU39607;  
 DT 19-JUN-2003 (first entry)  
 DE Protein encoded by Prokaryotic essential gene #25134.  
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 OS Pseudomonas putida.  
 PN WO200277183-A2.  
 PD 03-OCT-2002.  
 PF 21-MAR-2002; 2002WO-US009107.  
 PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 PA (ELIT-) ELITRA PHARM INC.  
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GT, Yamamoto R, Forsyth RA, Xu HH;  
 PI WPI; 2003-029926/02.  
 DR N-PSDB; ACA43477.  
 PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 PS Claim 25; SEQ ID NO 67531; 1766pp; English.  
 XX





QY 96 NNPDMILSNKQGVGYC---VDLF-----AGDLKGLKDKIPYFQELGLTYLHMLPFK 145  
Db 11 NDPQWYKD-----AVIYQVHVKSFYDANNDGIGDFAGLIEKLDYIADLGVNTLWLLPFY- 64  
QY 146 CPEGKSDGYAVSSYRDVNPALGTIGDLREVIALHHEAGISAVVDFIFNHTSNEHEW--- 202  
Db 65 -PSPRRDDGYDIAQYRGVHSDYGLADARRFIAEAHRGRLRVITELVINHTSDQHPWFI 123  
QY 203 AQRCAAGDPLFDNFYIIFPDRMPDQYDRTLREIFPDQHPGFSQLEDGRWWT----- 256  
Db 124 ARHAKKGSRRAD--YYWSD--SDEKYQGT-RIIFID-----TEQSNWTWDPVAQOY 170  
QY 257 ---TNSFQWDLNYSNPWFBRAMAGEMFLANLGVDILRMDAVAFIWKMGTSCEMLPQA 313  
Db 171 YWHRFYSHQPDINFDPQVLEVLGVMRXYLDMGVGDGLRDAIPYLIERDGTSSENLPET 230  
QY 314 HALIRAFNAVRIAPAVPFKSEAIVHPDQVVQYIG-----QDECQIGYNPLQWALLWN 367  
Db 231 HQVLKRIRAEIDAHYPDRMLLAEANQWPEPTRPYFGGEGEGECHMAFHFPILPRMYM 290  
QY 368 TLATREVNLLHQALTYRHNLEHTAWVNVYRSHDDIGWTFADEDAAYLGISGYDHRQFLN 427  
Db 291 AIAQEDRYPTIDILRQTPDIPANCQWAIFLRNHDELTEWVTD-----DERDYL- 339  
QY 428 RFEVNRFDGSFARGVPFQVNPSTGDCRVS---GTAALYGLAQDDPHAVDRIKLYSIAL 484  
Db 340 -----NMHYAADRRARLNLGIRRRRLAPVERDRR---RIELHSLLL 378  
QY 485 STGGLPLIYLGEVGT-----LNDDD-----WSQDSN-----KSDDSRWAHRPRYNEAL 528  
Db 379 SMPGTPTLYYGDIEGMGDNITLYGDRDGVRTPMQMSVDRNGGFSRADPAKLV-LPPIIDPL 437  
QY 529 Y-----AQRNDPSTAAGQIYQGLRHMIAVRQSNPRPDGRLVTFNTNKHITGYI-- 578  
Db 438 YGYQTINVEAQARDPHS---LINWMRLLAVRSQOKAFGRGSLKMLAPSNRRILAYIRE 493  
QY 579 -----RNNALLAFGNFSEYPTVTAHTLQAMPEKAH-----DLIGKTV---SLNQDL 623  
Db 494 YAEGERQDSTLCVANTLSRAAQAV-----ELDLASHAGKVPVEMIGMSFPPIGELTYLL 547  
QY 624 TLQPYQVMMLEIA 636  
Db 548 TLPPYGFYWEYLA 560

RESULT 14  
ABU38432  
ID ABU38432 standard; protein; 1100 AA.  
XX  
AC ABU38432;  
XX  
DT 19-JUN-2003 (first entry)  
XX  
DE Protein encoded by Prokaryotic essential gene #23959.  
XX  
KM Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX  
OS Pseudomonas aeruginosa.  
XX  
PN WO200277183-A2.  
XX  
PD 03-OCT-2002.  
XX  
PF 21-MAR-2002; 2002WO-US009107.  
XX  
PR 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zysekind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX  
XX WPI; 2003-029926/02.  
DR N-PSDB; ACA42302.  
DR  
XX  
XX  
PT New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
PS  
PS Claim 25; SEQ ID NO 66356; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 1100 AA;

Query Match 15.9%; Score 541.5; DB 6; Length 1100;  
Best Local Similarity 26.9%; Pred. No. 6.4e-43;  
Matches 165; Conservative 90; Mismatches 223; Indels 135; Gaps 25;

QY 96 NNPDMILSNKQGVGYC---VDLF-----AGDLKGLKDKIPYFQELGLTYLHMLPFK 145  
Db 11 NDPQWYKD-----AVIYQVHVKSFYDANNDGIGDFAGLIEKLDYIADLGVNTLWLLPFY- 64  
QY 146 CPEGKSDGYAVSSYRDVNPALGTIGDLREVIALHHEAGISAVVDFIFNHTSNEHEW--- 202  
Db 65 -PSPRRDDGYDIAQYRGVHSDYGLADARRFIAEAHRGRLRVITELVINHTSDQHPWFI 123  
QY 203 AQRCAAGDPLFDNFYIIFPDRMPDQYDRTLREIFPDQHPGFSQLEDGRWWT----- 256  
Db 124 ARHAKKGSRRAD--YYWSD--SDEKYQGT-RIIFID-----TEQSNWTWDPVAQOY 170  
QY 257 ---TNSFQWDLNYSNPWFBRAMAGEMFLANLGVDILRMDAVAFIWKMGTSCEMLPQA 313  
Db 171 YWHRFYSHQPDINFDPQVLEVLGVMRXYLDMGVGDGLRDAIPYLIERDGTSSENLPET 230  
QY 314 HALIRAFNAVRIAPAVPFKSEAIVHPDQVVQYIG-----QDECQIGYNPLQWALLWN 367  
Db 231 HQVLKRIRAEIDAHYPDRMLLAEANQWPEPTRPYFGGEGEGECHMAFHFPILPRMYM 290  
QY 368 TLATREVNLLHQALTYRHNLEHTAWVNVYRSHDDIGWTFADEDAAYLGISGYDHRQFLN 427  
Db 291 AIAQEDRYPTIDILRQTPDIPANCQWAIFLRNHDELTEWVTD-----DERDYL- 339



QY 428 RFFVNRFDGSPARGVPOYNPSTGDCRV---GTAALVGLAODDPHVAVDRIKILYSIAL 484  
Db 340 -----WNHYAADRRLRLNIGIRRLAPLVERDR---RIELHSLIL 378  
QY 485 STGLPLIYGDVEVT-----LNDDD-----WSQDSN---KSDSRWAHRPRYNEAL 528  
Db 379 SMPGTPLIYGYDEIGMDNITLGRDGVRTPMQMSVDRNGGFSRADPAKL-V-LPILDP 437  
QY 529 Y-----AQNDPSTAGQIYQGLRHMIAVRQSNPRFDGRLVTENTNKHITGYI-- 578  
Db 438 YGYQTINVEAQRDPHS---LLNWMRRLLAVRSQOKAFGRGSLKMLAPSNRILAYLR 493  
QY 579 -----RNALLAFGNFSEYPTVTAHTLQAMPFKAH-----DLIGKTV-----SLND 623  
Db 494 YAEGERQDSLICVANLSRAQAV-----ELDLASHAGKVPVEMIGMSFPPIGELTYL 547  
QY 624 TLQPYQVMWLEIA 636  
Db 548 TLPPYGFYWFYLA 560

RESULT 15

ABU21832  
ID ABU21832 standard; protein; 1154 AA.

AC ABU21832;

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #7359.

KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Burkholderia fungorum.

PN WO200277183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

PA (ELIT-) ELITRA PHARM INC.

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KU, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX WPI; 2003-029926/02.  
DR N-PSDB; ACA25702.

PT New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 49756; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation and the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

CC SQ Sequence 1154 AA;

Query Match 15.9%; Score 541; DB 6; Length 1154;  
Best Local Similarity 27.6%; Pred. No. 7.8e-43;  
Matches 155; Conservative 86; Mismatches 224; Indels 96; Gaps 18;

QY 118 GDLKGLDKIPYFQELGLTYLHMLPLFKCEPKSGDGYAVSSYRDVNPALGTIGDLREVI 177  
Db 65 GDFPGLIAKLDYIAELGVNAIMLPLFY--PSPRRDDGYDIADYRNHVPDYGNLSVKRFI 122  
QY 178 AALHRAGISAVVDFTFNHTSNEHEWAQRCAGDPLFDN-FYITFPPDRMPDQYDRTLREI 236  
Db 123 QEAHARGIRVITELVINHTSDQHPFQARARAKPGSNHRYVWSN---TDQKYLETRII 179  
QY 237 FPDQHPGFSQLED--GRWVTTTNSFQWDLNYSNPWFPMAGEMLFLANLGVDIRM 293  
Db 180 FIDSEPSNWT--HDPVAGAYYHRYFYSHPDPLNFDNPAVLREVLQVRFWLDMGIDGLRL 237  
QY 294 DAVAFWKQMGTSCEMLPOAHALIRAFNAVMRIAAPAVFPKSEAIVHPDQVQYIG-QDE 352  
Db 238 DAVPYLVEREENTNENLPETHAVLKRIRATIDAEYPMRLIAEANQWPEVKEYFGDEDE 297  
QY 353 CQIGYNPLQMALMNTLATREVNLLHQALTYRHNLPEHTAWNVYRSHDDIGWTF-ADED 411  
Db 298 CHMAFHPFLPRIYMSIASDEPFITDIMRQTPDLAETNQWAI FLRNHDELTLLEMTDSE 357  
QY 412 AAYL-GISGYDHRQFLNRFVNRFDGSPARGVPOYNPSTGDCRVSGTAALVGLAODDP 470  
Db 358 RDYLMNTYASDRRLN-----LGIRRLAPLIMERDR 389  
QY 471 HAVDRIKLILYSIALSTGLPLIYGDVEVT-----LNDDD-----WSQDSN---KSD 515  
Db 390 R---RIELINSLLSMPTGVITYGDELGMGDNHILGDRDGVRTPMQWSSDRNGGFSRAD 446  
QY 516 SRWAHRPRYNEALY-----AQNDPSTAGQIYQGLRHMIAVRQSNPRFDGRLVT 567  
Db 447 PEGLVLPVMGSLYGFDAVNVESAQRDPHS---LLNWMTRRLATRAKQTFGRGTIRFL 502  
QY 568 NTNKHITGYIR---NNALLAFGNFSEYPTVTAHTLQAMPFKAHDLIGKTVSLND- 622  
Db 503 KPENRKILAYLREMGEPIICVANLSRAP-----QAVELDLSEFDSVPITEMTADS 554  
QY 623 -----LTQPYQVMW 632  
Db 555 VPPPIGQLTYLTLTPPYGFLW 575

Search completed: May 25, 2004, 19:30:26  
Job time : 66 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 19:22:06 ; Search time 45 Seconds  
(without alignments)  
4459.324 Million cell updates/sec

Title: US-09-807-146-1

Perfect score: 3408  
Sequence: 1 MLPTQQVGLILQYLKTRIL.....VSLNQDLTLQPYQVWMLLEIA 636

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3408	100.0	636	2 Q9ZEU2	Q9zeu2 neisseria p
2	3338	97.9	636	2 Q84HD6	Q84hd6 neisseria m
3	3333	97.8	636	2 Q84HD5	Q84hd5 neisseria p
4	1646	48.3	701	16 Q7UG17	Q7ug17 rhodospirillum rubrum
5	1319.5	38.7	644	16 Q9RVT9	Q9rvt9 deinococcus
6	1126.5	33.1	644	16 Q8PGX2	Q8pgx2 xanthomonas
7	1116.5	32.8	637	16 Q8P512	Q8p512 xanthomonas
8	919.5	27.0	584	16 Q9A959	Q9a959 caulobacter
9	550	16.1	1106	16 Q88FN0	Q88fn0 pseudomonas
10	548	16.1	1160	16 Q8XT75	Q8xt75 ralstonia s
11	541.5	15.9	1100	16 Q91IW3	Q91iw3 pseudomonas
12	541.5	15.9	1113	16 Q7WIT8	Q7wit8 bordetella
13	538.5	15.8	1113	16 Q7VYJ9	Q7vyj9 bordetella
14	529.5	15.5	572	16 O54207	O54207 streptomyces
15	529.5	15.5	572	16 Q82UF2	Q82uf2 streptomyces
16	529.5	15.5	598	16 Q8NNM1	Q8nnm1 corynebacterium

17	529	15.5	1116	16 Q8PE49	Q8pe49 xanthomonas
18	525.5	15.4	1108	16 Q8XAR7	Q8xar7 chlorobium
19	523.5	15.4	612	16 Q8FND9	Q8fnd9 corynebacterium
20	522	15.3	1098	16 Q89FD4	Q89fd4 bradyrhizobium
21	522	15.3	1116	16 Q8PR14	Q8pr14 xanthomonas
22	521.5	15.3	1108	16 Q881X1	Q881x1 pseudomonas
23	520.5	15.3	552	16 Q9RST7	Q9rst7 deinococcus
24	520	15.3	965	2 Q7WU15	Q7wu15 thermus the
25	517.5	15.2	668	17 Q8TQ88	Q8tqa8 methanobacterium
26	516	15.1	566	2 O54203	O54203 streptomyces
27	515.5	15.1	566	16 Q91IK3	Q91ik3 streptomyces
28	515	15.1	965	2 Q9RA59	Q9ra59 thermus cal
29	511.5	15.0	583	16 Q82SR1	Q82sr1 streptomyces
30	509.5	15.0	601	16 Q07176	Q07176 mycobacterium
31	509.5	15.0	601	16 Q7U2S8	Q7u2s8 mycobacterium
32	484	14.2	585	2 Q7WUM4	Q7wum4 actinoplanes
33	467	13.7	553	16 Q7U4V8	Q7u4v8 synecococcus
34	452.5	13.3	568	16 Q8FHS2	Q8fhs2 escherichia
35	450.5	13.2	571	16 Q7UIS9	Q7uis9 rhodospirillum
36	445	13.1	544	16 Q92X63	Q92x63 rhizobium m
37	404.5	11.9	585	16 Q7V1B8	Q7vlb8 prochlorococcus
38	404	11.9	565	16 Q89VZ1	Q89vz1 bradyrhizobium
39	393	11.5	566	16 Q926R8	Q926r8 listeria in
40	386.5	11.3	434	16 Q8X816	Q8x816 escherichia
41	373	10.9	560	16 Q8Y3U6	Q8y3u6 listeria mo
42	372	10.9	547	16 Q88ZP5	Q88zp5 lactobacillus
43	349	10.2	557	16 Q88ZX0	Q88zx0 lactobacillus
44	345.5	10.1	575	16 Q7U3J7	Q7u3j7 synecococcus
45	342	10.0	537	16 Q835M8	Q835m8 enterococcus

#### ALIGNMENTS

##### RESULT 1

ID	Q9ZEU2	PRELIMINARY;	PRT;	636 AA.
AC	Q9ZEU2;			
DT	01-MAY-1999 (TrEMBLrel. 10, Created)			
DT	01-MAY-1999 (TrEMBLrel. 10, Last sequence update)			
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)			
DE	Amylosucrase (EC 2.4.1.4) (Fragment).			
GN	AMYLOSUCRASE.			
OS	Neisseria polysaccharaea.			
OC	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;			
OC	Neisseriaceae; Neisseria.			
OX	NCBI_TaxID=489;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 43768;			
RX	MEDLINE=99102197; PubMed=9882648;			
RA	Potocki de Montalk G., Remaud-Simeon M., Willemot R.M., Planhot V.,			
RA	Monsan P.;			
RT	"Sequence analysis of the gene encoding amylosucrase from Neisseria			
RT	polysaccharaea and characterization of the recombinant enzyme.";			
RL	J. Bacteriol. 181:375-381(1999).			
DR	EMBL; AJ011781; CAA09772.1; ..			
DR	PDB; 1G5A; 07-NOV-01.			
DR	PDB; 1JG9; 31-OCT-01.			
DR	PDB; 1JGI; 31-OCT-01.			
DR	PDB; 1MVY; 18-DEC-02.			
DR	PDB; 1MW0; 18-DEC-02.			
DR	PDB; 1MW1; 18-DEC-02.			
DR	PDB; 1MW2; 18-DEC-02.			
DR	PDB; 1MW3; 18-DEC-02.			
DR	GO; GO:0004556; F:alpha-amylase activity; IEA.			
DR	GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.			
DR	GO; GO:0005975; P:carbohydrate metabolism; IEA.			
DR	InterPro; IPR006047; Alpha-amyl_cat.			
DR	Pfam; PF00128; alpha-amylase; 1.			
KW	Glycosyltransferase; Transferase.			
FT	NON_TER			
SQ	SEQUENCE 636 AA; 72343 MW; B7656C19BFA2065 CRC64;			

Query Match	100.0%;	Score 3408;	DB 2;	Length 636;
Best Local Similarity	100.0%;	Pred. No. 1.7e-265;		
Matches 636;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	MLPTQOVGLILQYLKTRILDIYTPEDRAGIEKSEDMRQFSRRMDTHFPKLMNELDSVYG	60
Db	1	MLPTQOVGLILQYLKTRILDIYTPEDRAGIEKSEDMRQFSRRMDTHFPKLMNELDSVYG	60
QY	61	NNEALLPMLEMLLAQAWOSYSQRNSSLKIDIDARENPNPDWILSNKQGVGCYVDLPAAGDL	120
Db	61	NNEALLPMLEMLLAQAWOSYSQRNSSLKIDIDARENPNPDWILSNKQGVGCYVDLPAAGDL	120
QY	121	KGLKDKIPYFOELGLTYLHLMLPLFKCBGKSDGGAVSSYRDVNPALGTIGDLREVTAAL	180
Db	121	KGLKDKIPYFOELGLTYLHLMLPLFKCBGKSDGGAVSSYRDVNPALGTIGDLREVTAAL	180
QY	181	HEAGISAVVDPIFNHTSNEHEWAQRCAAGDPLFDNFYIIPDDRMPDQYDRITREIFPDQ	240
Db	181	HEAGISAVVDPIFNHTSNEHEWAQRCAAGDPLFDNFYIIPDDRMPDQYDRITREIFPDQ	240
QY	241	HPGGFSQLEDGRWVWTTFNSFQWDLNYSNPVFRAMAGEMFLANLGVDIRMDAVAFAIW	300
Db	241	HPGGFSQLEDGRWVWTTFNSFQWDLNYSNPVFRAMAGEMFLANLGVDIRMDAVAFAIW	300
QY	301	KQMGTSCENLPOAHALIRAFNAVMRIAAPAVEFKSEALVHPDQVQYIGODECOIGYNPL	360
Db	301	KQMGTSCENLPOAHALIRAFNAVMRIAAPAVEFKSEALVHPDQVQYIGODECOIGYNPL	360
QY	361	QMALMNTLATREVNLLHQALTYRHNLPBHTAWNVYVRSHDDIGWTFADBDAAVLGISGY	420
Db	361	QMALMNTLATREVNLLHQALTYRHNLPBHTAWNVYVRSHDDIGWTFADBDAAVLGISGY	420
QY	421	DHRQFLNREFVNRFDGSGFARGVPFQYNPSTGDCRVSGTAALVGLAODDPHAYDRIKLKY	480
Db	421	DHRQFLNREFVNRFDGSGFARGVPFQYNPSTGDCRVSGTAALVGLAODDPHAYDRIKLKY	480
QY	481	STATSTGGLPLIYLGDVEGTLNDDWGSQDSNKSDDSRWAHRPRYNEALYAQRNDBSTAAG	540
Db	481	STATSTGGLPLIYLGDVEGTLNDDWGSQDSNKSDDSRWAHRPRYNEALYAQRNDBSTAAG	540
QY	541	QIYQGLRHMIAYRQSNPRFDGGRLVTFNTNNKHIIIGYIRNNALLAFGNFSEYPTVTAHT	600
Db	541	QIYQGLRHMIAYRQSNPRFDGGRLVTFNTNNKHIIIGYIRNNALLAFGNFSEYPTVTAHT	600
QY	601	LQAMPKADHLIGKTVSLNODLTLPQYQVMLEIA	636
Db	601	LQAMPKADHLIGKTVSLNODLTLPQYQVMLEIA	636
RESULT 2			
ID	Q84HD6	PRELIMINARY;	PRT; 636 AA.
AC	Q84HD6;		
DT	01-JUN-2003	(Tremblrel. 24, Created)	
DT	01-JUN-2003	(Tremblrel. 24, Last sequence update)	
DT	01-OCT-2003	(Tremblrel. 25, Last annotation update)	
DE	Amylostrase Ams.		
GN	AMS.		
OS	Neisseria meningitidis.		
OC	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;		
OX	Neisseriaceae; Neisseria.		
NC	NCBI_TaxID=487;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=93246;		
RX	PubMed=12517860;		
RA	Zhu P., Tsang R.S.W., Tsai C.M.;		
RT	"Nonencapsulated Neisseria meningitidis Strain Produces Amylopectin		
RT	from Sucrose: Altering the Concept for Differentiation between N.		
RT	meningitidis and N. polysacchara."		
RL	J. Clin. Microbiol. 41:273-278 (2003).		
DR	EMBL; AY099334; AAM51152.1; -.		

DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro; IPR006047; Alpha\_amyl\_cat.  
DR InterPro; IPR006589; Alp\_amyl\_cat\_sub.  
DR Pfam; PF00128; alpha-amylase; 1.  
DR SMART; SM00642; Amyy; 1.  
SQ SEQUENCE 636 AA; 72339 MW; 4EE1D80F1F55B37D CRC64;

Query Match	97.9%	Score 3338;	DB 2;	length 636;
Best Local Similarity	97.8%	Pred. No. 7.5e-260;		
Matches 622; Conservative	7;	Mismatches 7;	Indels 0;	Gaps 0;

QY	1	MLPTQOVGLILQYLKTRILDIYTPQAGIEKSEDMRQFSRRMDTHPKLMNELDSVYG	60
Db	1	MLPTQOVGLILQYLKTRILDIYTPQAGIEKSEDMRQFSRRMDTHPKLMNELDSVYG	60
QY	61	NNEALLPMLMLLAQAWQSYSQNSSLKDDIDARENNPDWILSNKQGVGYVDLFAAGDL	120
Db	61	NNEALLPMLMLLAQAWQSYSQNSSLKDDIDARENNPDWILSNKQGVGYVDLFAAGDL	120
QY	121	KGLKDKIPYFOELGLTYLHMLPLFKCPEKSGSDGYAVSSYRDVNPALGTIGDLREVIAL	180
Db	121	KGLKDKIHYPQELGLTYLHMLPLFKCPEKSGSDGYAVSSYRDVNPALGTIGDLREVIAL	180
QY	181	HEAGISAVVDPIFNHTSNHEHMAQRCAAGDPLFDNFYITFPDRMPDQYDRTLREITFPDQ	240
Db	181	HEAGISAVVDPIFNHTSNHEHMAQRCAAGDPLFDNFYITFPDRMPDQYDRTLREITFPDQ	240
QY	241	HPGGFSQLEDGRWWTTFNSFQMDLINSNPWVFRAMAGEMFLANLGVDIRMDAVAFIW	300
Db	241	HPGGFSQLEDGRWWTTFNSFQMDLINSNPWVFRAMAGEMFLANLGVDIRMDAVAFIW	300
QY	301	KQMGTSCEMLPOAHALIRAFNAVMRIAAPAVFEKSEAIVHPDQVVQYIGQDECOIGYNPL	360
Db	301	KQMGTSCEMLPOAHALIRAFNAVMRIAAPAVFEKSEAIVHPDQVVQYIGQDECOIGYNPL	360
QY	361	QMALIMNTLATREVNLLHQALTYRHNLPEHTAMVNVYRSHDDIGWTFADEDAAYLGISGY	420
Db	361	QMALIMNTLATREVNLLHQALTYRHNLPEHTAMVNVYRSHDDIGWTFADEDAAYLGISGY	420
QY	421	DHRQFLNRFVNRFDGSPARGVPFOYNPSTGDCRVSGTAALVGLAQDDPHAVDRIKLXY	480
Db	421	DHRQFLNRFVNRFDGSPARGVPFOYNPSTGDCRVSGTAALVGLAQDDPHAVDRIKLXY	480
QY	481	SIALSTGGLPLIYLGDVEGTINDDDWSQDSNKSDDSRWAHRPRYNEALYAQRNDPSTAAG	540
Db	481	SIALSTGGLPLIYLGDVEGTINDDDWSQDSNKSDDSRWAHRPRYNEALYAQRNDPSTAAG	540
QY	541	QIYQGLRHMIAVRQSNPRFDGRLVTENTNNKHIIIGYIRNNALLAFGNFSEYQVTAAHT	600
Db	541	QIYQGLRHMIAVRQSNPRFDGRLVTENTNNKHIIIGYIRNNALLAFGNFSEYQVTAAHT	600
QY	601	LQAMPFKAHDLLIGKTVSLNQDLTLQPYQVMWLEIA	636
Db	601	LQAMPFKAHDLLIGKTVSLNQDLTLQPYQVMWLEIA	636

RESULT 3

ID	Q84HDS	PRELIMINARY;	FRT;	636	AA.
AC	Q84HDS;				
DT	01-JUN-2003	(TREMBLrel. 24, Created)			
DT	01-JUN-2003	(TREMBLrel. 24, Last sequence update)			
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)			
DE	Amylosucrase Ams.				
GN	AMS.				
OS	Neisseria polysaccharea.				
OC	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;				
OC	Neisseriaceae; Neisseria.				
OX	NCBI_TaxID=489;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=85322;				

RX PubMed=12517860;  
RA Zhu P., Tsang R.S.W., Tsai C.M.;  
RT "Nonencapsulated Neisseria meningitidis Strain Produces Amylopectin  
from Sucrose: Altering the Concept for Differentiation between N.  
meningitidis and N. polysacchara.";  
RL J. Clin. Microbiol. 41:273-278(2003).  
DR EMBL; AY099335; AAM51153.1; -  
DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro; IPR006047; Alpha\_amyl\_cat.  
DR InterPro; IPR006589; Alp\_amyl\_cat\_sub.  
DR Pfam; PF00128; alpha-amylase; 1.  
DR SMART; SM00642; Aamy; 1.  
SQ SEQUENCE 636 AA; 72271 MW; 98F81E0BFAC39A09 CRC64;

Query Match 97.8%; Score 3333; DB 2; Length 636;  
Best Local Similarity 97.5%; Pred. No. 1.9e-259;  
Matches 620; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

```
QY 1 MLPTQVGLILQYLKTRILDIYTPQAGIEKSEDMRQFSRRMDTHFKLMNELDSVYG 60
D 1 MLPTQVGLILQHLKTRILDIYTPQAGIEKSEDMRQFSRRMDTHFKLMNELDSVYG 60
QY 61 NNEALLPMLMLLAQAMQSYQSSSLKDIDIAKENNPDWILSNKQVGVGYVDFAGDL 120
D 61 NNEALLPMLMLLAQAMQSYQSSSLKDIDIAKENNPDWILSNKQVGVGYVDFAGDL 120
QY 121 KGLKDKIPYQELGLTYLHMLPFCPEGKSDGGYAVSSYRDVNPALGTIGDLREVIAL 180
D 121 KGLKDKIPYQELGLTYLHMLPFCPEGKSDGGYAVSSYRDVNPALGTIGDLREVIAL 180
QY 181 HEAGISAVVDYFIENHTSNEHEWAQCAAGDPLFDNFYIIFPDRMPDQYDRTLREIFPD 240
D 181 HEAGISAVVDYFIENHTSNEHEWAQCAAGDPLFDNFYIIFPDRMPDQYDRTLREIFPD 240
QY 241 HPGGFSQLEDGRWWTTFNSFQMDLNYSNPWFRAMAGEMFLANLGVDIRMDAVAFIW 300
D 241 HPGGFSQLEDGRWWTTFNSFQMDLNYSNPWFRAMAGEMFLANLGVDIRMDAVAFIW 300
QY 301 KQMGTCENLPQAHALIRAFNAVMRIAAPVFFKSEAIHPDQVQYIGQDECOIGYNPL 360
D 301 KQMGTCENLPQAHALIRAFNAVMRIAAPVFFKSEAIHPDQVQYIGQDECOIGYNPL 360
QY 361 QMALMNTLATREVNLLHQALTYRHNLPBHTAVNYSRSHDIGNTPADEDAAYLGISGY 420
D 361 QMALMNTLATREVNLLHQALTYRHNLPBHTAVNYSRSHDIGNTPADEDAAYLGISGY 420
QY 421 DHRQFLNRFVNRFDGSFARGVPFQYNPSTGDCRVSGTAALVGLAQDDPHAVDRIKLY 480
D 421 DHRQFLNRFVNRFDGSFARGVPFQYNPSTGDCRVSGTAALVGLAQDDPHAVDRIKLY 480
QY 481 SIALSTGGLPLIYLGEVGTINDDWSQDSNKSDDSRWAHRPRYNEALYAQRNDPSTAAG 540
D 481 SIALSTGGLPLIYLGEVGTINDDWSQDSNKSDDSRWAHRPRYNEALYAQRNDPSTAAG 540
QY 541 QIYQGLRHMIAVROSNPFRDGRVTFNTNNKHIIIGYIRNNALLAFGNFSEYPTVTAHT 600
D 541 QIYQGLRHMIAVROSNPFRDGRVTFNTNNKHIIIGYIRNNALLAFGNFSEYPTVTAHT 600
QY 601 LQAMPFRKADLIGKTVSLNODLTLOPYQVMWLEIA 636
D 601 LQAMPFRKADLIGKTVSLNODLTLOPYQVMWLEIA 636
QY 636 LQAMPFRKADLIGKTVSLNODLTLOPYQVMWLEIA 636
D 636 LQAMPFRKADLIGKTVSLNODLTLOPYQVMWLEIA 636
```

## RESULT 4

Q7UGI7 PRELIMINARY; PRT; 701 AA.  
AC Q7UGI7;  
DT 01-OCT-2003 (Tremblrel. 25, Created)  
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Alpha-amylase, amylousucrase (EC 2.4.1.4).  
GN RB5196.

OS Rhodopirellula baltica.  
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;  
OC Planctomycetaceae; Pirellula.  
OX NCBI\_TaxID=117;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1;  
RX MEDLINE=22735913; PubMed=12835416;  
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,  
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,  
RA Schlesner H., Aumann R., Reinhardt R.;  
RT "Complete genome sequence of the marine planctomycete Pirellula sp.  
strain 1.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).  
DR EMBL; BX294141; CAD78342.1; -  
KW Glycosyltransferase; Transferase; Complete proteome.  
SQ SEQUENCE 701 AA; 81070 MW; 29FF2FD85C45F454 CRC64;

Query Match 48.3%; Score 1646; DB 16; Length 701;  
Best Local Similarity 52.1%; Pred. No. 1.5e-123;  
Matches 316; Conservative 94; Mismatches 180; Indels 16; Gaps 6;

```
QY 43 RMDTHFKLMNELDSVYGNNEALLPMLMLLAQAMQSYQSSSLKDIDIAKENNPDWIL 102
D 94 RIYQHWDDLFGYLLHLYGDQWDFEYHLEQILLTIRGWRMRPEQLKQDDEHRINNPEWYQ 153
QY 103 SNKQVGVGYVDFAGDLKGLKDKIPYQELGLTYLHMLPFCPEGKSDGGYAVSSYRD 162
D 154 SEKLVGALYVDLFSNLGELRQKIPYQDGLSYLHMLPFAVRPQNNNDGYAISNYS 213
QY 163 VNPALGTIGDLREVIALHEAGISAVVDYFIENHTSNEHEWAQCAAGDPLFDNFYIIFPD 222
D 214 VDPRTGIDDLRLIADDLRAGILLVLDVENHTADDHWAQQAQSGNBEYQKYFIIFPD 273
QY 223 RRPMDQYDRTLREIFPDQHPGFSQLEDG--RWVWTTFNSFQMDLNYSNPWFRAMAGEM 280
D 274 REVPDQYERTLREIFPTVRGNFT-WHDGMOQWWTTFNSFQMDLNYRNPEVFRAMLSEM 332
QY 281 LFLANLGVDIRMDAVAFIWKMGTCENLPQAHALIRAFNAVMRIAAPVFFKSEAIHV 340
D 333 LFLANTGVDIRLDAVAFIWKMGTCENLPQAHALIRAFNAVMRIAAPVFFKSEAIHV 392
QY 341 PDQVVOYIGQDECOIGYNPLQMALMNTLATREVNLLHQALTYRHNLPBHTAVNYSRSH 400
D 393 PDDVVRYSHKECQTSYNPTLMAIMWESLATRDTKLRRSLARRHKLPERTVWNYLRCH 452
QY 401 DDIGWTFADEDAAYLGISGYDRQFLNRFVNRFDGSFARGVPFQYNPSTGDCRVSGTAA 460
D 453 DDIGWTFDDADAAYGIDAFGRQFLNAFYTGQFEGSFARGVPFQHPNDGMRISGTILA 512
QY 461 ALVGLAQ-----DD--PHAVDRIKLYSIALSTGGLPLIYLGEVGTINDDWSQDSNK 512
D 513 SLAGEQAIESDRDDWKELAIRILLINAMMLSVGGIPLYLGEWGAALNDYDFVKPAK 572
QY 513 SDDSRWAHRPRYNEALYAQRNDPSTAAGQIYQGLRHMIAVROSNPFRDGRVTFNTNNK 572
D 573 AGDTRWVHRPKMQWKFLELGDGDSIRGRITRSIRKLIALRKMTPAFGGLDMELLVLES 632
QY 573 HIIGYIR--NNALLAFGNFSEYPTVTAHTLQ--AMPFRKADLIGKTVSLNODLTLOP 627
D 633 HILGYVRQHEGNRVIVIANPCEHAQEIDSNRLRTAGMGFRFEDLYSGQITTTSHPLTGP 692
QY 628 YQVMWL 633
D 693 YEFMWL 698
```

## RESULT 5

Q9RVT9 PRELIMINARY; PRT; 644 AA.  
AC Q9RVT9;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)



```
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Alpha-AMYLASE.
GN DR0933.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Usterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AE001946; AAF10510.1; -.
DR PIR; C75457; C75457.
DR TIGR; DR0933; -.
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha_amyl_cat.
DR Pfam; PF00128; alpha-amylase; 1.
KW Lyase; Complete proteome.
SQ SEQUENCE 644 AA; 71668 MW; 6100775611D81978 CRC64;

Query Match 38.7%; Score 1319.5; DB 16; Length 644;
Best Local Similarity 42.1%; Pred. No. 2.6e-97;
Matches 281; Conservative 110; Mismatches 215; Indels 61; Gaps 14;

QY 1 MLTPQGVGLILQYLKTRILDIYTPQAGIEKSEDMRQFSRMDTHPKLMNELDSVYG 60
DB 1 MLTPD-----LAARL-----RLAFDDDRDAETFRLEERYGPGLADNLRAYVG 43
QY 61 NN-EALL-PMLEMLLAQAWQSYQRNSLKDIDIAENNPDWILSNKQVGVYVDLFAG 118
DB 4 NHADALIGELLEVM-----HAYHARPADLRUDEARLLRPDWLQPEMWGYVAYVDRFAG 99
QY 119 DLKGLKDKIPYFOELGLTYLHMLPKCEBEGKSDGYAVSSYRDVNPALGTIGDLREVI 178
DB 100 TLRGVGERLEYLEGLVTTYLHLLPLRPDGENGGYAVQDYRSVRPDLGTIDLSALAR 159
QY 179 ALHEAGISAVVDPIFNHTSNEHEWAQRCAAGDPLFDNFYIFPDRMPDQYDRTREIF? 238
DB 160 ELRGRGISLVLDVLNHNVAEBHEWAVRATGEAAVRYDFHIFPRTQPDAYERLPEIF? 219
QY 239 DQHPGFGS-QLEDGRVWVTTFNSFQWDINYSNPVFRAMAGEMFLANLGVDIRMDAVA 297
DB 220 DFAPGNFTWNGEAGGWVWTTFNRSQMDVWNGNPAVFRREYLDLILTLANRGEVEFRIDAIA 279
QY 298 FIMKQMGTSCEMLPQAHALIRAFNAVRIAPAVEFKSEATVHPDQVQYIGQDE----- 352
DB 280 FLWKRLGTDCQNPQEVHRLTHALRAATRIYAPAVAFKABAIYAPGDLIHYLSRDHGRV 339
QY 353 CQIGYNPLQWALLNLTATREVNLIHQALTYRHNLPEHTAVWVYVRSHDQIGWTFADEDA 412
DB 340 SDMAYHNSLWQMLSSLASRDTLLTALAAPFPKPTNTTWGVYVRCHDDIGWAIADEDA 399
QY 413 AYLGISGVDRQFLNRFVNRFDGSPARQVFPQYNPSTGDCRVSGTAAALVGL----- 465
DB 400 ARVGLSGPAHRHFLSDFYSGEPGSGFARGLVFQHPQTDGRRISGTIASLAGDLALETG 459
QY 466 -AQDDPHAVDRIKLKYSIALSTGGLPLIYGDEVGTINDDDWSQDSKSDSRWAHRPRY 524
DB 460 DAERVNDALRLLLHVAVMLGEGVPLTYMGDELALINDTFAAVPAHAADNRWVRPQ 519
QY 525 NEALYAAQ-----RNDPSTAAGQIYQGLRHMIAVRQSNPRFDGGRLVTENTNKH----- 574
```

```
DB 520 DWELVASAQADAATGQPTPAGRMFAGLRHLAVRRTPHILHA-----STESRPLPSD 573
QY 575 --IGYIR---NNALAFGNFSEYPQTVTAHTLQ-AMPPKAHDLIGKTVSL-NQDLTLQ 626
DB 574 PCVLLIRREHPTGVLLQVYNFSEHITPTWPLQEQUGAVAHDLIGESQFHLGPPDLAE 633
QY 627 PYQVMWL 633
DB 634 PYRALMWL 640

RESULT 6
Q8PGX2 PRELIMINARY; PRT; 644 AA.
ID Q8PGX2;
AC Q8PGX2;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Amylosucrase or alpha amylase.
GN XAC3490.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarote G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA localis E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AE011997; AAM38333.1; -.
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha_amyl_cat.
DR Pfam; PF00128; alpha-amylase; 1.
KW Complete proteome.
SQ SEQUENCE 644 AA; 70470 MW; 041ED026FE935CCD CRC64;

Query Match 33.1%; Score 1126.5; DB 16; Length 644;
Best Local Similarity 39.0%; Pred. No. 9e-82;
Matches 240; Conservative 107; Mismatches 236; Indels 33; Gaps 10;

QY 43 RMDTHPKLMNELDSVYGNNEALLPMLEMLLAQAWQSYQRNSLKDIDIAENNPDWIL 102
DB 30 RYDQHASRLLDALHALYQGRADYASWLAQWLGBVGDIAQRQPALQTLDDSTR--HAGWFG 87
QY 103 SNKQVGVYVDLFAAGDLKGLKDKIPYFOELGLTYLHMLPKCEBEGKSDGYAVSSYRD 162
DB 88 QQHMLGYSAYADRFAAGTLQGVAAERYPYLQELGVRYLHLLPLRARAQDNDGFAVSDYQ 147
QY 163 VNPALGTIGDLREVIAPALHEAGISAVVDPIFNHTSNEHEWAQRCAAGDPLFDNFYIFPD 222
DB 148 VEPSTGNDIDLVALTSRLREAGISLCADFVLNHTADDAWMAQARAQADARYLDYHHFAD 207
QY 223 RRMPPDQYDRTLREIFPDQHPGFSQLED-GRVWVTTFNSFQWDINYSNPVFRAMAGEML 281
```

```

Db      208  RTLPDRYEATLGQVFPHTAPGNFTWDDTAQOMMTTPEYPYQWDLNWSNPAVFGMALAML 267
QY      282  FLANLGVDLRMDAVAFIWKOMGTSCENLFOAHALIRAFNAVMRIAAPAVEPKSEAIYHP 341
Db      268  RLANLGVETERLDSTAYLWKRIGTDCNQOBEAHTLLVALRAVTDIVAPAVVMKAEAIYPM 327
QY      342  DQVYQYIGQ-----DECQIGYNPLQMALMLMTLATREVNLLHQALTYRHNLPBHTAWVNY 396
Db      328  TQLPPEYFGSGVDEGHECHLAHSTLMAAGSALALQRGDILHNVIASHPLPRHCAMLSTY 387
QY      397  VRSHDDIGWTFADEDAAYLGISGYDHRQFLNRFVNRFPDGSFARGVPFOYNPSTGD--CR 454
Db      388  VRCHDDIGMNVNLQHEACGNAAQPPFSLRDVARFYANAVPGSYARGESFQ--SSGDGVHG 444
QY      455  VSGTAALVGL--AODDPH-----AVDRIKLLYSIALSTGGLPLYLGDEVGTINDDDW 506
Db      445  TNGMAAALAGIQAAQOAGDAALAAVAADVLLLYALATALAMPGPVPLYMGDELA MVNDPGY 504
QY      507  SQDSNKSDDSRWahrPRYNEALYAQRNDSPTAAGQIYQGLRHMIAVRQSNPRFDGR-LV 565
Db      505  RHPDRQHEGRWLHRPAMDWQLAQRHDANSLSGTVYRRLRGLIRQRAALGALAADQALA 564
QY      566  TFNTNNKHIIIGYIRNNALLAFGNFSEYPTV-----TAHTLQAMPFKAHDLIGKTVS 618
Db      565  SIALNDPRVFLVLTGRGDSFIALHNFSQPLDVELLAIGIDGWTLLAI---DDAIGGAAR 620
QY      619  LNQDITLQPYQVMWLE 634
Db      621  GDGSIPLPPYGVRLQ 636

```

RESULT 7	
08P512	
ID	Q8P512 PRELIMINARY; PRT; 637 AA.
AC	Q8P512;
DT	01-OCT-2002 (TREMBLrel. 22, Created)
DT	01-OCT-2002 (TREMBLrel. 22, last sequence update)
DT	01-JUN-2003 (TREMBLrel. 24, last annotation update)
DE	Amylasecrase or alpha amylase.
GN	XCC3359.
OS	Xanthomonas campestris (pv. campestris).
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC	Xanthomonadaceae; Xanthomonas.
OX	NCBI_TaxID=340;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=ATCC 33913 / NCPPB 528;
RC	MEDLINE=22022145; PubMed=12024217;
RA	da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA	Queiroz R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA	Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA	Camarote G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA	Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA	Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA	Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA	Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA	Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA	Martins E.C., Melandris J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA	Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA	Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA	Spinoia L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA	Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA	Setubal J.C., Kitajima J.P.;
RT	"Comparison of the genomes of two Xanthomonas pathogens with differing
RT	host specificities.";
RL	Nature 417:459-463 (2002).
DR	EMBL; AB012454; AAM42629.1; -
DR	GO; GO:0004556; F:alpha-amylase activity; IEA.
DR	GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR	InterPro; IPR006047; Alpha_amyl_cat.
DR	Pfam; PF00128; alpha-amylase; 1.
KW	Complete proteome.
SA	SEQUENCE 637 AA; 69295 MW; B42C5BF7B94ADFAE CRC64;

Query Match	32.8%;	Score 1116.5;	DB 16;	Length 637;
Best Local Similarity	39.3%;	Pred. No. 5.6e-81;		
Matches 241;	Conservative 106;	Mismatches 236;	Indels 31;	Gaps 10
QY	43	RMDTHFPKLMNELDSVYGNNEALLPMLLEMLLAQAMQSYSGQNSSLKDIIDARENPPDWIL	102	
Db	31	RFDTHAPRLLDALSTLYGSHADYVTWLPQWLSALGVAAQARPAALRHLDLSRA--PGWFG	88	
QY	103	SNKQVGVCYVDLFAGDLKGLKDKIPYFOELGLTYLHMLPLFKCEGKSDGAYVSSYRD	162	
Db	89	QQDMLGYSAYVDRFAGTLRGVAERVPYQLQELGVRYLHLLPFLRARGDNDGFAVSDYGQ	148	
QY	163	VNPALGTIGDLREVIALHEAGISAVVDPIFNHTSNEHEWAQRCAADPLFDNFYIIFPD	222	
Db	149	VEPALGSNDLDVALTARLRANISLCADFVLNHTADHDHAWAQARAAGDTRYLDYHHFAD	208	
QY	223	RRMPDQYDRTLREIFPDQHPGFSQLEDGR-VWTTTFNSFQWDLNYSNPWFRAMAGEML	281	
Db	209	RNAPDQYDRTLTVQVFPQTAPGNFTWDETRQMMWTFYPYQWDLNWSNPVAFEGEMALAML	268	
QY	282	FLANLQVDILRMDAVAFIWKQMGTSCENLPOAHALIRAFNAVMTIAAPVFEKSEATVHP	341	
Db	269	ELANLQVEAFRLDSTAYLWKRPETNCMNQPEAHTIILVALRAVADIVAPSVMKABAIVPM	328	
QY	342	DQVQYIIGQ----DECQIGVNPLOMALLNNTLATREVNLIHQALTYRHNLPEHTAWNY	396	
Db	329	AELPPYFGSGVQRGHECHLAHYHSTILMAAGWSALALQRGDILQDVIAHSPPLPNCAMLSY	388	
QY	397	VRSHDDIGWTFADEDAAYLGISGDHRQFLNRFVFNRFDSFARGVPFQYNPSTGD--CR	454	
Db	389	VRCHDDIGWNVLOHEAAGTAQDPFSLREVAQFYANAVPGSYARGESFQ--SSGDGVHG	445	
QY	455	VSGTAALVGLAQDDPH-----AVDRIKLILYSIALSTGGLPLIYLGDVEGTLNDDDW	506	
Db	446	TNGMSAALVGVQAHEHADAAAARAVIDRLVLLYAVSLAMPGVPLIYMGDELLPNDTAY	505	
QY	507	SQDSNKSDDSRWAHRPRYNEALYAQRNDPSTAAGQIYQGLRHMIAVRQSNPRFDGGR-LV	565	
Db	506	LDDAQQRHREGRWLHRPAMAWELLAQOHDASTLAGVTYTRLALIRLRAGLPALAAATQSLG	565	
QY	566	TFTNTNKHIIIGYIRNNALLAFGNFSEYP---QVTAAHTLQAMPFKAHDLIG-GKTVSLN	620	
Db	566	SVALGDARLFPALTRGDSFLAVHNFSVDVPLPVDDLQTGHALWAV---LDTDGTGDABEPH	621	
QY	621	QDLTLQPYQVMWLE 634		
Db	622	TELLIPAYGVRWLQ 635		

RESULT 8		
Q9A959		
ID Q9A959	PRELIMINARY;	PRT; 584 AA.
AC Q9A959;		
DT 01-JUN-2001 (TREMBLrel. 17, Created)		
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)		
DE Amylosucrase.		
GN CCL135.		
OS Caulobacter crescentus.		
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;		
OC Caulobacteraceae; Caulobacter.		
OX NCBI_TaxID=155892;		
OX [1]		
RE SEQUENCE FROM N.A.		
RC STRAIN=ATCC 19089 / CB15;		
RA MEDLINE=21173698; PubMed=11259647;		
RA Niernan W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,		
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohca N., Maddock J.R.,		
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,		
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,		
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,		
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,		

RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
RT "Complete genome sequence of *Caulobacter crescentus*,"  
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141 (2001).  
DR EMBL; AE005791; AAK23119.1; --  
DR PIR; C87390; C87390.  
DR TIGR; CC1135; --  
DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro; IPR006047; Alpha amyl cat.  
DR Pfam; PF00128; alpha-amylase; 1.  
KW Complete proteome.  
SQ SEQUENCE 584 AA; 63232 MW; D2788FE875ACD624 CRC64;

Query Match 27.0%; Score 919.5; DB 16; Length 584;  
Best local Similarity 35.2%; Pred. No. 3.6e-65;  
Matches 215; Conservative 100; Mismatches 246; Indels 49; Gaps 16;

QY 44 MDTHFPKLMNEDSVYGNNEALLPMLMLAQAWQSYSORNSLKDIDIDARENNPDWILS 103  
DB 1 MESRFAKL-----YGADARGPAVLERLKTNLLKAHARPEPLRALDARAADPAWLHA 53  
QY 104 NKQGVGVYVDLFAGDLKGLKDKIPYFOELGLTYLHLMPLFKCPEKSGDGYAVSSYRDV 163  
DB 54 PGQTAVTFYVDRFAGDLNGVRGKLDYLTGLVRLHPLPLEPRPGSDGFAVADYRKV 113  
QY 164 NPALGTIGLREVIATLHEAGISAVVDLIFNHTSNEHEWAQRCAAGDPLFDNFYIFPDR 223  
DB 114 DPLGTIDLEALAGDLRQRDMGLLDVVCNHTAREHAWAKARAGDPAYRDYIVLPDA 173  
QY 224 RMPDQYRTRLEIFPDQHPGGS-QLEDGRWWTTFNSFQWDLNYSNPWFRAMAGEMLF 282  
DB 174 QSAARDELIDVFPDTPAGSFITYDAAMGQVWTFYFPQWDLNYPVFAEMLEVLIF 233  
QY 283 LANLGVDLRMDAVAFWKQMGTSCEMLPQAHALIRAFNAVRIADAVFKESEIVHPD 342  
DB 234 LAAKGAQGRLDASAPFLWKQAGTTCRNLPQTYEIVEAWRAALSIAPGVLLAEALIESVE 293  
QY 343 QVVQYIGODE--CQIGYNPLQMALMTTLATREVNLLHQALTYRHLPEHTAMVNVRSI 400  
DB 294 DVLPFEGEESGCMLAYNNVMTALMAALADDAVIARCLAVAARKPAQGAWLNVRCR 353  
QY 401 DDIGWTFADEDAAYLGISGYDHRQFLNRFVNRFDGSFARGVPFQYN----PSTGDCRVS 456  
DB 354 DDLIMNAL--AAVAPAS--DLRWSNAY--GNNEG-FSRGRAFGTAEGVPSST-----N 400  
QY 457 GTAAALVGLAQD--DPHAVDRIKLYSIALSTGGLPLIYLGDEVGTLDDWSDSNKS 513  
DB 401 GMAAALAGLTADADGCLGARRRLLYGITHALDGMPLIYMGDEIGLDNDEAYQDDPLRA 460  
QY 514 DDSRWahrPRYNEALYAQRNDPSTAAGQY-----QGLRHMIAVRQSNPRPDGRLVT 566  
DB 461 GDRWLHRPQMDWSIAERRGEAGALQADLFAFARLQGRARLATLGYAGP---ARPV- 515  
QY 567 FNTNNKHIIIGYIRNNA---LIAFGNFSEYPTVTAHTLQAMPFKAHDLIGKTVSLNQL 623  
DB 516 -EVSSPAVALFLRDEGARPFICVANVSDAPQDF--ELPPAFAQGAEDVDLG-APSPAGAV 571  
QY 624 TLQPYQVMWL 633  
DB 572 SLPPYGITWL 581

RESULT 9  
Q88FNO PRELIMINARY; PRT; 1106 AA.  
ID Q88FNO  
AC Q88FNO; 01-JUN-2003 (Tremblrel. 24, Created)  
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Trehalose synthase, putative.  
GN PPA059.  
OS Pseudomonas putida (strain KT2440).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=160488;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22423060; PubMed=12534463;  
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,  
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,  
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,  
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,  
RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzez A.,  
RA Utecherback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,  
RA Lauber J., Stjepandic D., Hohnsels J., Straetz M., Heim S.,  
RA Kiewitz C., Eissen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,  
RA Fraser C.M.;  
RT "Complete genome sequence and comparative analysis of the  
RT metabolically versatile *Pseudomonas putida* KT2440.";  
RL Environ. Microbiol. 4:799-808 (2002).  
DR EMBL; AE016789; AAN69649.1; --  
DR TIGR; PPA059; --  
DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0003910; F:DNA ligase (ATP) activity; IEA.  
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR GO; GO:0006310; P:DNA recombination; IEA.  
DR GO; GO:0006281; P:DNA repair; IEA.  
DR GO; GO:0006260; P:DNA replication; IEA.  
DR InterPro; IPR006047; Alpha amyl cat.  
DR InterPro; IPR000977; DNA ligase.  
DR Pfam; PF00128; alpha-amylase; 1.  
DR PROSITE; PS00697; DNA\_LIGASE\_A1; 1.  
KW Complete proteome.  
SQ SEQUENCE 1106 AA; 125589 MW; 2F5E471637096B38 CRC64;

Query Match 16.1%; Score 550; DB 16; Length 1106;  
Best local Similarity 29.0%; Pred. No. 5.5e-35;  
Matches 164; Conservative 83; Mismatches 228; Indels 90; Gaps 20;

QY 118 GDLKGLDKIPYFOELGLTYLHLMPLFKCPEKSGDGYAVSSYRDVNPALGTIGLREVI 177  
DB 39 GDFAGLSKLDYIALGVNTLWLPPY--PSPRRDDGYDAEYKAVHPDYGSMDARRFI 96  
QY 178 AALHEAGISAVVDLIFNHTSNEHEWAQR--CAAGDPLFDNFYIFPDRMPDQ-YDRTL 233  
DB 97 AEAHKGRGLVITELVINHTSDQHPQARHAKGSKARE--FYWSD--DDQKYDGT- 150  
QY 234 REIFPDQHPGGS-QLEDGRWWTTFNSFQWDLNYSNPWFRAMAGEMFLANLGVDIR 292  
DB 151 RIIFLDEKSNWMTDPAQYFMRHFSHQPDILNFPQVLKAVGWRFWLDLGVGDLR 210  
QY 293 MDAVAFWKQMGTSCEMLPQAHALIRAFNAVRIADAVFKESEIVHPDQVQYIGQ-- 350  
DB 211 LDAIPYLIERDGTNNENLAETHDVAKAIRAEIDANYPDRMLLAENQWPEPTRPYFGEED 270  
QY 351 -DECGIGNPLQMALMTTLATREVNLLHQALTYRHLPEHTAMVNVRSRDDIGWTF-A 408  
DB 271 GDECHMAHFPLMPRMYMALAMEDRPFITDILRQTPETPANCQWALFLRNHDELTEMVT 330  
QY 409 DEDAAVYLGISGYDHRQFLNRFVNRFDGSFARGVPFQYNPSTGDCRVS-GTAAALVGLAQ 467  
DB 331 DRERDYL-----WNYYAEDRRARINLGI RRRRLAPLLO 362  
QY 468 DDPHAVDRIKLYSIALSTGGLPLIYLGDEVGT-----LNDSD-----WSQDSN----- 511  
DB 363 RDRR---RIELLTSLLSMPGTFTLYGDELGMGDNITYLGDRDGVTRPMQWSPDRNGGFS 419  
QY 512 KSDSRWAHRPRYNEALY-----AQRNDPSTAAGQIYQGLRHMIAVRQSNPRPDGGR 563  
DB 420 RADPQRLV-LPIMDPLYGYQTVNVEAQSHDPS---LNNWTRMLAVRKQOKAFGRGS 474  
QY 564 LVTFTNNKHIIIGYR-----NNALLAFGNFSEYPTVTAHTLQAMPFKAHDLIGK 615  
DB 475 LRTLTFSNRILAYIRBYTDADGNTEVILCVANVSRAAQAALIELISQYADKVPVEMLGGS 534



QY 616 TV-----SLNODLTLPYQVMMLEIA 636  
Db 535 AFPPIGQLPELTLTPYAFYWFLLA 559

RESULT 10

Q8XT75 PRELIMINARY; PRT; 1160 AA.

AC Q8XT75;  
DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Hypothetical protein RSP0240.  
GN RSP0240 OR RS05183.  
OS Ralstonia solanacearum (Pseudomonas solanacearum).  
OC Plasmid megaplasmid.  
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
OC Burkholderiaceae; Ralstonia.  
OX NCBI\_TaxID=305;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GMT1000;  
RX MEDLINE=21681879; PubMed=11823852;  
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
RA Arlat M., Billault A., Brottier P., Camus J.C., Catolico L.,  
RA Chandler M., Choise N., Cladel-Renard C., Cumnac S., Demange N.,  
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,  
RA Sigulier P., Thebaud P., Whalen M., Wincker P., Levy M.,  
RA Weissenbach J., Boucher C.A.;  
RT "Genome sequence of the plant pathogen Ralstonia solanacearum."  
RL Nature 415:497-502(2002).  
DR EMBL; AL646077; CAD17391.1; -  
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.  
DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro; IPR006047; Alpha\_amyl\_cat.  
DR InterPro; IPR006589; Alp\_amyl\_cat\_sub.  
DR Pfam; PF00128; alpha-amylase; 1.  
DR SMART; SM00642; Amy; 1.  
KW Plasmid; Hypothetical protein; Complete proteome.  
SQ SEQUENCE 1160 AA; 129108 MW; 27555FA3ACEBDE55 CRC64;

Query Match 16.1%; Score 548; DB 16; Length 1160;  
Best Local Similarity 29.3%; Pred. No. 8.5e-35;  
Matches 164; Conservative 79; Mismatches 229; Indels 88; Gaps 20;

QY 118 GDLKGLKDKIPYFQELGLTYLHMLPFCPEKSGDGYAVSSYRDVNPALGTIGDLREVI 177  
Db 38 GDFPGLISKLDYIAELGVDAIWLPPFY--PSPRDDGYDIAEYRGVHPDYGTMDARRFI 95  
QY 178 AALHEAGISAVVDPIFNHTSNEHEW---AQRCAAGDPLFDNFYITFDRRMPDQYDRTLR 234  
Db 96 AEAHARGLRVITTELVINHTSDQHPWFQARRAKAGSALRD--FYWSDH---DKKYAGTR 150  
QY 235 EIFPDQHPGFS--QLEDGRWVWTTFNSFQMDINYSNPWFRAMAGEMFLANLGVDIR 292  
Db 151 IIFIDTEPSNWTWDPVADA--YVWHRFYSHQPDINFDPRLKAVLGVKFWLNLGVGDLR 209  
QY 293 MDAVAFIWKQMGTSCENLPOAHALIRAFNAVMRIAAPVFFKSEALVHPDQVQYIGQ-D 351  
Db 210 LDVAPYLVEREGTANENLPETHAVLRKIRAMDAEFKNRLLLAEANQWPEDTQYFNGND 269  
QY 352 ECQIGYNPLQWALLMNTLATRBNVNLHQAALTYRHNLPEHTAVNVVYRSHDDIGWTF-ADE 410  
Db 270 ECHMAFHPLMRMYMAIAREDRPITDIHQTPPEVPTCQWALFENHDELTEMVTDA 329  
QY 411 DAAYL-GISGYDHRQFLNRFVNRFDGSFARGVPFQYNPSTGDCRVSGTAALVGLAODD 469  
Db 330 ERDYLMEVYASDRRARIN-----LGIRRLAPLLERD 361  
QY 470 PHAVDRIKLYSIALSTGGLPIYLGEVGT-----LNDD-----WSQDSN---KSD 514  
Db 362 RR--RVELMNSILFSMPGTVMYIGDEIGMGDNIHLGDRDGVRTPMQWSPDRNGGFSRA 418

QY 515 DSRWAHRPRYNEALY-----AQRNDPSTAQGIYQGLRHMIAVROSNPFRDGRVT 566  
Db 419 DPEQLVLPAIMGSLGYRVSVEAQTRD---AHSILNWTIRLLATRKHRVFGRGSIQF 474

QY 567 FNTNKKHIIIGYIR-----NNALLAFGNFSEYPTV---TAHTLQAMPFKAHDLIGKTV-- 617  
Db 475 LQPNRKVLAYARALEGEAPILCVANTLSRASQAVELDLAAYAQRVPV---ELIGGTAFPP 531

QY 618 --SLNODLTLPYQVMMLEI 635

Db 532 IGQLPYLTLTPYAFYWFLEL 551

RESULT 11

Q91IW3 PRELIMINARY; PRT; 1100 AA.

ID Q91IW3  
AC Q91IW3;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Probable trehalose synthase.  
GN PA2152.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 15692 / PAOI;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,  
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Ladbis K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PAOI, an  
RT opportunistic pathogen."  
RL Nature 406:959-964(2000).  
DR EMBL; AE004642; AAG05540.1; --  
DR PIR; G83376; G83376.  
DR HSSP; P21332; IUOK.  
DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro; IPR006047; Alpha\_amyl\_cat.  
DR InterPro; IPR006589; Alp\_amyl\_cat\_sub.  
DR Pfam; PF00128; alpha-amylase; 1.  
DR SMART; SM00642; Amy; 1.  
KW Complete proteome.  
SQ SEQUENCE 1100 AA; 124521 MW; 00BE2B8CE8D8A8AF CRC64;

Query Match 15.9%; Score 541.5; DB 16; Length 1100;  
Best Local Similarity 26.9%; Pred. No. 2.6e-34;  
Matches 165; Conservative 90; Mismatches 223; Indels 135; Gaps 25;

QY 96 NNPDWISNKQVGVY---VDLF-----AGDLKGLKDKIPYFQELGLTYLHMLPFLK 145  
Db 11 NDPQWYKD-----AVIYQVHVKSFYDANNNDGIGDFAGLIEKLDYIADLGVNTLWLLPFY- 64  
QY 146 CPEKSGDGYAVSSYRDVNPALGTIGDLREVI AALHEAGISAVVDPIFNHTSNEHEW--- 202  
Db 65 -PSPRDDGYDIAQYRGVHSDYGLADARRFIABAHRRGLRVITTELVINHTSDQHPFIR 123  
QY 203 AQRCAAGDPLFDNFYITFDRRMPDQYDRTLREIFPDQHPGFSQLEDGRWVWT----- 256  
Db 124 ARHAKKGSARD--YVWSD--SDEKQGT-RIFID-----TEQSNWTWDPVAQQY 170  
QY 257 ---TFNSFQMDINYSNPWFRAMAGEMFLANLGVDIRMDAVAFIWKQMGTSCENLPOA 313  
Db 171 YVWHRFYSHQPDINFDPQVLRREV.LGVWERYWLDMGVDGLRLDAIPYLIERDGTSSENLPET 230  
QY 314 HALIRAFNAVMRIAAPVFFKSEALVHPDQVQYIG-----QDEQIGYNPLQWALLWN 367



```
DB 231 HOVLKRIRAELEDAHYPDRLMLAEANQWPEOTRPPYFGGEDGEGDECHMAFHPPLMPRMV 290
QY 368 TLATREVNLIHQALTYRHNLPEHTAMVNVYRSHDDIGWTFADEDAAYLGISGYDHRQFLN 427
DB 291 AIAQEDRYPITDILRQTPDI PANQWAI FLRNHDEL TLEMVTD-----DERDYL- 339
QY 428 RFEVNRFDGSFARGVPEQYNPSTGDCRSV---GTAALVGLAODDEPHAVDRIKLYSTIA- 484
DB 340 -----WNHYAADRRARLNLGIRRLAPLVERDR-RIELHSLT- 378
QY 485 STGGLPLIYLGEVGT-----LNDD-----WSQDSN---KSDSRWAHRPRYNEA- 528
DB 379 SMPGTPILYGEDEIGMGDNIVLGD RDGVRTPMQWSVDRNGGFSRADPAKL V-LPPLDP- 437
QY 529 Y-----AQENDPSTAAGIYQGLRHMLAVRQSNPRFDGRLVTENTNKHIIIGYI- 578
DB 438 YGYQTINVEAQARDPHS-----LNNMRRLLAVRSQOKAFGRGLKMLAPSNRRILAYLRE 493
QY 579 -----RNNALLAFGNFSEYPTQVTAHTLQAMPFKAH-----DLIGKTV---SLNOD- 623
DB 494 YAEGERQDSILCVANLSRAQAV-----ELDLASHAGKVPEMIGMSFPPIGELTYL- 547
QY 624 TLQPYQVWMLLEIA 636
DB 548 TLPPYGFYFYLA 560
```

## RESULT 12

```
Q7WI18 PRELIMINARY; PRT; 1113 AA.
AC Q7WI18;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Alpha-glucosidase.
GN BB2863.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RA "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640445; CAE33355.1; -.
KW Complete proteome.
SQ SEQUENCE 1113 AA; 124163 MW; AFG52A4B054940D1 CRC64;
```

Query Match 15.9%; Score 541.5; DB 16; Length 1113;  
Best local Similarity 27.2%; Pred. No. 2.7e-34;  
Matches 154; Conservative 90; Mismatches 223; Indels 99; Gaps 17;

```
QY 118 GDLKGLKDKIPYFQELGLTYLHMLPLFKCPGKSDGGA VSSYRDVNPALGTIGDLREVI 177
DB 35 GDFAGLIAKLDIYELGVNTIWL LPPY--DEPRRDDGYDIADYRGVHPDYGSLADARRLV 92
QY 178 AALHEAGISAVVDIFENHTSNEHEWAQRCAGDP-LFDNFYIIFPDRMPDQYDRTLREI 236
DB 93 RAAHARGRLVITELVINHTSDQHPWFQBARARPGSAHRAYYWMSD---DDKAYAGTRII 149
```

```
QY 237 FPDQHPGFS-QLEDGRWVWTFNSFQWDLNYSNPWVRAMAGEMLFLANLGVLDLRMDA 295
DB 150 FCDTERKSNWTDVPVAGAYFWHRPYSHQPDNLNDNPQVLRREVIGAMRYWLDMGVDGLRLDA 209
QY 296 VAFIWKQGTSCENLPQAHALIRAFNAVMRIAAPAVFEKSAIVHPQVQYIGQ-DECO 354
DB 210 VPVLVEREGTNENLPETHALIRIRRVIDSEYPGRMLLAEANQWPEDAOEYFGAGDECH 269
QY 355 IGVNPLQMALMNTLATREVNLIHQALTYRHNLPEHTAMVNVYRSHDDIG----- 404
DB 270 MAHFEPMLPRMYMAIAQEDRLPYTDIIRQTPSIAPQCQWAFILRNHDEL TLEMVTSRERD 329
QY 405 --W-TFADEDAAYLGISGYDHRQFLNREFVNRFDGSFARGVPEQYNPSTGDCRVSGTAAA 461
DB 330 YLMNVYAABPRARINL-----GIRRR 350
QY 462 LVGLAODDEPHAVDRIKLYSTIALSTGGLPLIYLGEVGT-----LNDD-----WSQD 509
DB 351 LAPLBERDR-RIELMNSLLSMPGTPVLYGDELGMGDNILHGD RDGVRTPMQWSPD 407
QY 510 SN---KSDSRWAHRPRYNEALY-----AQENDPSTAAGIYQGLRHMLAVRQSNP 557
DB 408 RNGGFSRADPER-LPLPLMGPLYGYEAVNVEAQORDPHS-----LNNTRMLAKRRQSH 462
QY 558 RFDGRLVTENTNKHIIIGYIR--NNALLAFGNFSEYPTQVTAHTLQAMPFKADLIGG 614
DB 463 VFGREHLSFLYPGNRKITAYLRWEDTVL CVANLSQAQPYEHLHSEYAGRVPEMLGG 522
QY 615 KTV---SLNODLTQPYQVWMLLEIA 636
DB 523 TAFPQIGELPYLLTLPPFGFYWLDS 548
```

## RESULT 13

```
Q7VYJ9 PRELIMINARY; PRT; 1113 AA.
AC Q7VYJ9;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Alpha-glucosidase.
GN BP1329.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama 1 / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RA "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640415; CAE41622.1; -.
KW Complete proteome.
SQ SEQUENCE 1113 AA; 124135 MW; 7235BDC938E49756 CRC64;
```

Query Match 15.8%; Score 538.5; DB 16; Length 1113;  
Best local Similarity 27.2%; Pred. No. 4.7e-34;  
Matches 154; Conservative 90; Mismatches 223; Indels 99; Gaps 17;

```
QY 118 GDLKGLKDKIPYFQELGLTYLHMLPLFKCPGKSDGGA VSSYRDVNPALGTIGDLREVI 177
```

```
Db 35 GDFAGLLAKLDYIVELGVNTIWLIPFY--PSPRDDGDVADIYRGVHPDYGSADARLLV 92
QY 178 AALHEAGISAVVDFIENHTSNEHEWAQRCAGDP-LFDNFYIIFPDRMPDQYDRLREI 236
Db 93 RAHAHARGLRVITELVINHTSDQHPWFQARARAPGSAHAYYVMSD---DDKAYACTRII 149
QY 237 FPDQHPGGFS-QLEDGRVWTTTNSFQWDLNYSNPVFRAMAGEMFLANLGVDIRMDA 295
Db 150 FCDTEKSNWTWDPVAGAYFVHRFYSHQPDNLNDNPQYLREVLGAMRYWLDGVDGLRIDA 209
QY 296 VAFIWKQMTSCENLPOAHALIRAFNAVRIAPAVFEKSEAIVHPDQVQYIGQ-DECO 354
Db 210 VPYLVREGTNNENLPEHTAILRIRRVIDSEYFGMLLAENQWPEDAQEYFGAGDECH 269
QY 355 IGYNPLQMALWNTLATREVNLLHQALTTRHNLPEHTAVVNVYRSHDDIG----- 404
Db 270 MAHFEPIMPRMYMAAQEDRLPVTDIIRQTPSIAPQOWAIFLRNHDELTEMVTSRERD 329
QY 405 --W-TFAEDDAAYLGISGYDHRQFLNRFVNRFDGSPFARGVPEQYNPSTGDCRYSGTAAA 461
Db 330 YLMNVVYAAEPARINL-----GIRRR 350
QY 462 LVGLAQDDPHAVDRIKLYSIALSTGGLPIYLGEVGT-----LNDDD-----WSQD 509
Db 351 LAPLLERDRR--RIELMNSLLISMPTPVLYGDELGMGDNHILGDRDGVRTPMQWSPD 407
QY 510 SN---KSDSRMAHPRRYNEALY-----AQRPSTAAQIYQGLRHMTAVRQSNP 557
Db 408 RGGFSRADPER-LPLPLMGPLYGEAVNVEAQQRDPHS---LNTWTRMLAKRRQSH 462
QY 558 RFDGRLVTFTNTNKHIIIGYIR--NNALLAFGNFSEYEQVTANTLQAMPFKADLIG 614
Db 463 VEGRGELSFYLPNGRKILAYLRTWEDTVLCVANLSQAQPVLEHLSEYAGRVEMLG 522
QY 615 KTV---SLNQDLTLQPYQVWMLAIA 636
Db 523 TAFPQIGELPYLTLPPFGFYWLDLS 548

RESULT 14
054207 PRELIMINARY; PRT; 572 AA.
ID 054207;
AC 054207;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
DE Putative trehalose synthase.
GN SC07334 OR SC4G10.13C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Schneider D., Bruton C.J., Chater K.F.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=96154943; PubMed=8596463;
RA Bruton C.J., Plaskitt K.A., Chater K.F.;
RT "Tissue-specific glycogen branching isoenzymes in a multicellular
RL prokaryote, Streptomyces coelicolor A3(2).";
RL Mol. Microbiol. 18:89-99(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=A3(2);
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RL the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabbinkwitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RL coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AJ001206; CAA04607.1; -.
DR EMBL; AL393131; CAB92880.1; -.
DR HSSP; P21332; 1UOK.
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha_amyl_cat.
DR InterPro; IPR006589; Alp_amyl_cat_sub.
DR InterPro; IPR008263; Glyco_hyd16_AS.
DR Pfam; PF00128; alpha-amylase; 1.
DR SMART; SM00642; Aamy; 1.
DR PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
DR Complete proteome.
KW SEQUENCE 572 AA; 65695 MW; 81252A982BBD51D CRC64;
SQ

Query Match 15.5%; Score 529.5; DB 16; Length 572;
Best Local Similarity 26.9%; Pred. No. 8.8e-34;
Matches 161; Conservative 85; Mismatches 244; Indels 109; Gaps 19;

QY 93 ARENPDWT-----LSNKQGVGYVYDLFAGDLKGLKXIPYQHLGLTYLHL 140
Db 16 AGDRHPDWFKRAVFEVLVRSFQDSNGDGI-----GDLKGLTAKLDYQLQWLGVDCLWL 68
QY 141 MPLFKCPGKSDGYAVSSYRDVNPALGTIGDLREVIALHEAGISAVVDFIENHTSNEH 200
Db 69 PPFKSP--LRDGYDVSDYTAVLPEFGDLADFEVDAHQGRVITDFVWHTSDQH 126
QY 201 EWAQ--RCAAGDPLFDNFYIIFPDRMPDQYDRLREIFPDQHPGGFSQLED--GRWW 255
Db 127 PMFQESRKNPDPGPDYVWADDTRRAD-----ARILFVDTASNWT--HDEVRGQYWW 179
QY 256 TTNSFQWDLNYSNPVFRAMAGEMFLANLGVDIRMDAVAFIWKQMTSCENLPOAHA 315
Db 180 HRFSSHQPDNLNENPAVQEBMALAKFWLDLGYDGYRLDAVPLYAEBGTNCENLPASHA 239
QY 316 LIRAFNAVRIAPAVFEKSEAIVHPDQVQYIGQ-----DECQIGYNPLQMALNNTLA 370
Db 240 FLKRVREIDAQYPTVLLAEANQWPEVDVDFGDIYSTGDCHEMAHFHPVRIEMAVR 299
QY 371 TREVNLLHQALTTRHNLPEHTAVVNVYRSHDDIGWTF-ADEDAAYLGISGYDHRQFLNRF 429
Db 300 RESRYPVSEILAKTPAIPSGCOWGIFLRNHDELTEMVTDSEEDYM----- 345
QY 430 FVNRFDGSPFARGVPEQYNPSTGDCRYSGTAAA 486
```

Db 346 -----YAEYAKDPRMRANIGIRRLATLLDNR---DQIELFTALLAL 386  
QY 487 GGLPLIYLGEVGT-----LNDDD-----WSQDSN---KSDSRWAHRPRYNEALY-- 529  
Db 387 PGSPILYYGDEIGMGDNIMLGDRDAVRTPMQWTPDRNAGFSTCDPGRLYLPALMDPVYGY 446  
QY 530 -----AQRNDPSTAAGQIYQGLRHMIAVRQSNPRFDGRLVTENTNKHIIIGYR--N 580  
Db 447 QVTNVEASMASPSS---LHWTTRMIEIRKQNPAGLGYTELPSNPAPVLAFLREYED 502  
QY 581 NALLAFGNFSEYPTVTAHTLQAMPFKAHDLIGKTV-----SLNQDLTLQPYQVMWLEI 635  
Db 503 DLVLCVNNFARFAQPTELDLREFAGRHVPVELFGVRFPAIGELPYLLTLGGHGFYMERL 561

RESULT 15

Q82JF2 PRELIMINARY; PRT; 572 AA.  
ID Q82JF2  
AC Q82JF2;  
DT 01-JUN-2003 (Tremblrel. 24, Created)  
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Putative trehalose synthase.  
GN SAV2803.  
OS Streptomyces avermitilis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Streptomycineae; Streptomycetaceae; Streptomyces.  
CX NCBI\_TaxID=33903;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
RX MEDLINE=21477403; PubMed=11572948;  
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,  
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;  
RT "Genome sequence of an industrial microorganism Streptomyces  
RT avermitilis: deducing the ability of producing secondary  
RT metabolites."  
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
RX MEDLINE=22608306; PubMed=12692562;  
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,  
RA Sakaki Y., Hattori M., Omura S.;  
RT "Complete genome sequence and comparative analysis of the industrial  
RT microorganism Streptomyces avermitilis."  
RL Nat. Biotechnol. 21:526-531(2003).  
DR EMBL; AP005032; BAC70514.1; .  
DR GO; GO:0004556; F:alpha-amylase activity; IEA.  
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro; IPR006047; Alpha\_amyl\_cat.  
DR InterPro; IPR006589; Alp\_amyl\_cat\_sub.  
DR Pfam; PF00128; alpha-amylase; 1.  
DR SMART; SM00642; Amy; 1.  
KW Complete proteome.  
SQ SEQUENCE 572 AA; 65751 MW; A6B0EAB49FF4D3EA CRC64;

Query Match 15.5%; Score 529.5; DB 16; Length 572;  
Best local similarity 27.2%; Pred. No. 8.8e-34;  
Matches 166; Conservative 86; Mismatches 233; Indels 125; Gaps 23;

QY 93 ARENNPDWT-----LSNKQVGVYVDLFAADLKGDKKIPYFOELGLTYLHL 140  
Db 16 AKDRDPEWFRKRAVFEVLVRSFQDSNGGV-----GDLKGLTAKLDYLGWLGVDCLWL 68  
QY 141 MPLFKCPGKSDGGYAVSSYRDVNPALGTIGDLREVIALHAEAGISAVVDPIFNHTSNEH 200  
Db 69 PPFKSP--LRDGYDVSDYTAVLPEFGDLADFEVDAHQGRMRYIIDFVMNHTSDLH 126  
QY 201 EWAQ--RCAAGDPLFDNFYIIFPDRMPDQYDRTLREIFPDQHPGHSQLEDGRWV---- 254  
Db 127 PMFQESRSNPDGPGYD--YYVWAD--DDKQYQDARIIFVD-----TEASNWTFDPV 173

QY 255 -----WTFENSFQMDINYSNPWFVRAMAGEMLELANLGVDILRMDAVAFIWKMGTSNEN 309  
Db 174 RKQYWMHRRFFSHQPDINYPENPAVQEEIVSALRFWLDIGIDGFRLDVAVPYLYQOEGTNCEN 233  
QY 310 LPQAHALIRAFNAWRIAAPAVFKESEAIYHPDQVYIGQ-----DECQIGYNPLQMAL 364  
Db 234 LPATHEFLKRVKKEIDTHYPDTVLLAEANQWPEVDYDFGDFPSGGDECHMAFHPVMPR 293  
QY 365 LWNLTATREVNLLHQALTYRHNLPEHTAMVNYVRSHDDIGWTF-ADEDAAYLGISGYDHR 423  
Db 294 IFMAVRRRSRYPSSELLAKTPAIPSSCQWGIFLRNHDELLTLEMVTDDEERDYMW----- 346  
QY 424 QFLNRFVNRFPDGSFARGVFPQYNPSTGDCRVSGTAAALVGLAODDPHAYDRIKLLYSIA 483  
Db 347 -----AEYAKDPRMRAN-----IGIRRLAPLLDNR---NQIELFTALL 383  
QY 484 LSTGGLPLIYLGEVGT-----LNDDD-----WSQDSN---KSDSRWAHRPRYNEAL 528  
Db 384 LSLPGSPILYYGDEIGMGDNIMLGDRDAVRTPMQWTPDRNAGFSSCDPGRLYLPALMDPV 443  
QY 529 Y-----AQRNDPSTAAGQIYQGLRHMIAVRQSNPRFDGRLVTENTNKHIIIGYR- 579  
Db 444 YGYQVTNVEASMSPPSS---LHWTTRMIEIRKQNPAGLGSYTELQSSNPAPVLAFLRE 499  
QY 580 -----NNALLAFGNFSEYPTVTAHTLQAMPFKAH--DLIGKTV---SLNQDLTL 625  
Db 500 APSTGNGDDLVLVCVHNSRFAQPTELDLRAFSGR-HVELLIGVRFPAIGELPYLLTL 557  
QY 626 QPYQVMWLEI 635  
Db 558 AGHGFYMERL 567

Search completed: May 25, 2004, 19:31:55  
Job time : 49 secs





```
OY 204 QRCAGDPLFDNFYFIIPDRRMPDQYDRTLREIFPDQHPGFS-QLEDGRWVWTTNSFQ 262
Db 124 SRSDPDGPGYGD--FYVWSD---TDELYQDARVIFVDTEPSNWTMDQTRGQYWHRRFHHQ 178
OY 263 WDLNYSNPWYERAMAGEMFLANLGLMDAVAFIWKQMGTSCEMLPOAHALIRAFNA 322
Db 179 PDLNFDNPKYQDAMLBAWFLMDGDFRLDAVPYLVERPGTNGENLPETHEMLKRYR 238
OY 323 VMRIAPAVFEKSEAIIVHPDQVQYIGDECOIG-----YNPLQMALNMTLAT 371
Db 239 FVDDNYFDRLVLYEANQMPDVBVYFGPEEREDGTVGPESHMAFHFPVMPRIF--MAV 295
OY 372 REVNLI--LHQALTYRHLPEHTAVVYVRSDDIGWTF-ADEDAAYLGISGYDHRQFLNR 428
Db 296 RRESRPISSEIMEQTPAIPGCGQGLFLRNHDELTEMVTDDEDRDYM----- 343
OY 429 FFVNRFDGSPARGVPFOYNPSTGDCRVSGTAALVGLAODDPAVDRIKLILYSIALSTGG 488
Db 344 -----GEYAKDPRMKAN-----IGIRRLAPLLDND---TNQIELFTALLSLPG 385
OY 489 LPLIYLGEVGTLDNDWSDSN-----KSDSRWA-----HRPRYNEALY--- 529
Db 386 SPVLYYGEIG-MGDNIWLGRDGVPTPMQRTPDNRNVGFSAAATPGKLHLPTIQDPVYGYQ 444
OY 530 -----AQRNDPSTAAGQIYQGLRHMTAVRQSNPRFDGRLVTFNTHKHIIGYIR----- 579
Db 445 SVNVEAQLBNPSS-----LHWTTRMTHIRQRDAFGLGTFEDLGSNPAVLSTVRELPGD 500
OY 580 --NNALLAFGNFSEYPTVTAHTLQAMPFKAH---DLIGG---KTVSLNODLTLQPYQV 630
Db 501 GGDVILCYNNLSRFPQPV---ELDLRYKRGVPELVIGVFPFPAVGELPYLLTSLGHGF 557
OY 631 MWLEI 635
Db 558 YWFERL 562

RESULT 2
TRES_THETH
ID_TRES_THETH STANDARD; PRT; 963 AA.
AC 006458;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Trehalose synthase (EC 5.4.99.16) (Maltose alpha-D-
glucosyltransferase).
GN TRES.
OS Thermus thermophilus.
OC Bacteria; Deinococcus-Thermus; Thermococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=274;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33923 / DSM 674 / AT-62;
RX MEDLINE=97194924; PubMed=9042362;
RA Tausaki K., Nishimoto T., Nakada T., Kubota M., Chaen H., Fukuda S.,
Sugimoto T., Kurimoto M.;
RT "Cloning and sequencing of trehalose synthase gene from Thermus
aerophilus ATCC33923."
RL Biochim. Biophys. Acta 1334:28-32(1997).
CC -1- FUNCTION: Catalyzes the conversion of maltose into alpha,alpha-
trehalose by transglucosylation.
CC -1- CATALYTIC ACTIVITY: Maltose = alpha,alpha-trehalose.
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL; D86216; BAAL9934.1; -.
DR HSSP; P21332; 1UOK.
DR InterPro; IPR006589; Alp_amy1_cat_sub.
DR InterPro; IPR006047; Alpha_amy1_cat.
DR Pfam; PF00128; alpha-amy1ase; 1.
DR SMART; SM00642; Amy; 1.
KW Isomerase.
SQ SEQUENCE 963 AA; 110170 MW; B6CB3CED365764AE CRC64;

Query Match 14.2%; Score 484; DB 1; length 963;
Best Local Similarity 26.4%; Pred. No. 2.8e-29;
Matches 146; Conservative 87; Mismatches 203; Indels 116; Gaps 22;

OY 118 GDLKGLDKIPYFOELGLTYLHIMPLFKCPGKSDGVAVSSYRDVNPALGTIGLREVI 177
Db 28 GDFEGLRKLPLYLEELGVNTLMLPFFQSP--LRDDGYDISDYQILPVHGTLEDL-TV 83
OY 178 AALHEAGISAVVDIFNHTSNEHWAQRC-AAGDPLFDNFYFIIPDRRMPDQYDRTLREI 236
Db 84 DEAHGRGMKVILVNLNHTSIDHPWFQEARKNPSMRD--WYVWSD--TPEKY-KGVYRVI 138
OY 237 FPDQHPGFSQLEDGRWV-----WTFNSFQWDLNYSNPWYERAMAGEMFLANLG 287
Db 139 PKD-----FETSNWTFDPAKAYYWHRFYWHQPDLMNDSPEVEKAIHQVMFFWADLG 190
OY 288 VDLRMDAVAFIWKQMGTSCEMLPOA-HALIRAFNAVRIAPAVFEKSEAIIVHPDQVQ 346
Db 191 VDGFRDLAIPLYEREGTSCENLPETIEAVKRLRKALBEERYGKILLAEVNMWPEETLP 250
OY 347 YIQQ-DECOIGYNPLQMALNMTLATREVNLLHQALTYRHLPEHTAVVYVRSDDIGW 405
Db 251 YFGDGVHMAYNFPLMPRIEMALRREDRGPIETMLKEABGIPETAQWALFLRNHDELTL 310
OY 406 TEADEDAAVYLGISGYDHRQFLNFFVNRFDGSPARGVPFOYNPSTGDCRVSGTAALVGL 465
Db 311 EKYTB-----EEREFMYE-----AYAPDPKFRIN-----LGIRRLMPL 344
OY 466 AQDDPHAVDRIKLILYSIALSTGGLPLIYLGEVGTLDND-----DWSQDSNK 512
Db 345 LGDDR---RYELLTALLTLKGTPIVYGDGIG-MGDNPFLGDRNGVRTPMQWSQDRIV 400
OY 513 SDDSRWAHR--PRYNEALY-----AQRNDPSTAAGQIYQGLRHMTAVRQSNPR-FD 560
Db 401 AFSRAPYHALFLPVSSEGYSTHFVNVEAQRENPHS---LLSFNRFLALRNQHAKEIFG 456
OY 561 GGRVLTFTNNKHIIIGYIRNN---ALLAFGNFSEYPTVTAHTLQAMPFKAHDLIGKTV 617
Db 457 RGSLLTLLPVENRVLAYLREHGERVLVVAANLSRYTQAF----- 495
OY 618 SLNODLTLQPYQ 629
Db 496 ---DPLPEAYQ 503

RESULT 3
SUCCP_ECOLI
ID_SUCCP_ECOLI STANDARD; PRT; 559 AA.
AC P76041; P78149; P78151;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative sucrose phosphorylase (EC 2.4.1.7) (Sucrose
glucosyltransferase).
GN YCJM OR B1309.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
```

RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12";  
RL Science 277:1453-1474 (1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12;  
RX MEDLINE=97251357; PubMed=9097039;  
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,  
RA Itch T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,  
RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,  
RA Motomura K., Nakade S., Nakamura Y., Nishimoto H., Nishio Y.,  
RA Oshima T., Saito N., Sempel G., Seki Y., Sivasundaram S.,  
RA Tagami H., Takeda J., Takekoto K., Takeuchi Y., Wada C.,  
RA Yamamoto Y., Horiuchi T.;  
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome  
corresponding to the 28.0-40.1 min region on the linkage map";  
RL DNA Res. 3:363-377 (1996).  
CC -1- CATALYTIC ACTIVITY: Sucrose + phosphate = D-fructose + alpha-D-  
glucose 1-phosphate.  
CC -1- SIMILARITY: Belongs to the sucrose phosphorylase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AE000229; AAC74391.1; ALT\_INIT.  
DR EMBL; D90768; BAA14878.1; -.  
DR EMBL; D90769; BAA14886.1; -.  
DR EcoGene; EGI3910; YCJM.  
DR InterPro; IPR006047; Alpha\_amyl\_cat.  
DR Pfam; PF00128; alpha-amylase; 1.  
KW Hypothetical protein; Transferase; Glycosyltransferase;  
KW Complete proteome.  
SQ SEQUENCE 559 AA; 64185 MW; C05B71B0E0E369C2 CRC64;  
  
Query Match 13.3%; Score 452.5; DB 1; Length 559;  
Best Local Similarity 26.7%; Pred. No. 3.5e-27;  
Matches 165; Conservative 98; Mismatches 260; Indels 95; Gaps 26;  
  
QY 50 KLMNELDSVYGNNEALLPMEMLIAQAMQSQRNSLKDIDIRENNPDWISNKQVGG 109  
DB 4 KITDYLDEIYGGT-----FTATHLQKLVTRLBSAKRL-ITQRKKHW--DESDVVL 51  
QY 110 VCYVDLF-AGDLKGLKDKIPYFOELG-----LTYLHLMPLFKCEPKSGDGYAVSSYR 161  
DB 52 ITYADQFHSNDLKPL-----PTNQFYHQWLQSFVHLLPFY---PSSSDGFSVIDYH 104  
QY 162 DVNPALGTIGDLREVTALHEAGISAVVDFFIFNHTSNEHEWAQRCAAGDPLFDPNFYITFP 221  
DB 105 QVASEAGEMODIQGLGESH-----LMPDFVCNHSMAKSEWFKNYLQHPGFEDFFI--- 156  
QY 222 DRMPDQYD-----RTLREIFPDQHPGSGFSLD--GRVWTTTNSFQWDLNYSNPWV 272  
DB 157 --AVDPQTDLSAVTRPRALPLTP-----FQMRDHSRHLMTTFSDDQIDNYSRSPV 207  
QY 273 FRAMAGEMFLANLGVLDILRMDAVAFIWKQMGTCENLPOAHALIRAFNAVRIAPAVF 332  
DB 208 LLAMVDVLLCYLAKGAEYVRIDAVGFMMKEPGTSCIHLEKTHLITKLRSITDNVAPGTV 267  
QY 333 EKSEA-IVHPDQVQY-IGQDECQIGYNPLQMALMNTLATREVNLL--HQALTYRHNL 387  
DB 268 ITTETNVPHKDNIAVFGAGDDEAHMVYQFSLPLVLAQKONVEALCAWAQNLTL---L 323  
QY 388 P-EHTAMVNVYRSHDDIGWT-----FADEDAAYIGISGYDHRQFLNRFVNFEDGSPARG 441  
DB 324 PSSNTTWENFLASHDGIQLNPLRGLLPESEILELVEALQOEGALVN--WKNNPDGTRS-- 379

QY 442 VFPQNPSTGDCRVSGTAAALVGLAODDPHAYDRIKLISIALSTGGLPLIYLGDEVGTL 501  
DB 380 -PYEINVTYMD-----ALSRESSDEERCARPLIAHALLSFGVPALYQSIGISR 430  
QY 502 NDDWWSQDSNKSQDSRWAHRPRY-NEALYAQRNDPSTAAGQIYQGLRHMAVROSNRPF 560  
DB 431 ND---YAGVEKLGYNRAINRKKYHSKEITRELNDEATLRNAVYHELRLITLRSHNEFH 487  
QY 561 GGRLVTEFNTNKHILGYIRNNA---LLAFGNFSEVPQTYTAHTLQAMPFKAHDLIGKT 616  
DB 488 PDNNFTIDTINSSVWRIPRSNADGNCITGLFNVSKNIQHVNITNLH-----GRDLISEVD 542  
QY 617 VSLNQDLTLQPYQVWMLE 634  
DB 543 I-LGNEITLRPWQVWMIK 559

RESULT 4  
MAL1 DROVI STANDARD; PRT; 586 AA.  
ID MAL1 DROVI  
AC 016098;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Maltase 1 precursor (EC 3.2.1.20).  
GN MAV1.  
OS Drosophila virilis (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidae; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7244;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=9;  
RX MEDLINE=97475684; PubMed=9335139;  
RA Vieira C.P., Vieira J., Hartl D.L.;  
RT "The evolution of small gene clusters: evidence for an independent  
RT origin of the maltase gene cluster in Drosophila virilis and  
RT Drosophila melanogaster.";  
RL Mol. Biol. Evol. 14:985-993 (1997).  
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing 1,4-  
linked D-glucose residues with release of D-glucose.  
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF006573; AAB82327.1; -.  
DR HSSP; P21332; IUOK.  
DR FlyBase; FBgn0022839; Dvir\MAV1.  
DR InterPro; IPR006589; Alp\_amyl\_cat\_sub.  
DR InterPro; IPR006047; Alpha\_amyl\_cat.  
DR Pfam; PF00128; alpha-amylase; 1.  
DR SMART; SM00642; Aamy; 1.  
KW Hydroxylase; Glycosidase; Signal.  
FT SIGNAL 19  
FT CHAIN 1  
FT CARBOHYD 20 586  
FT CARBOHYD 133 133  
FT CARBOHYD 166 166  
FT CARBOHYD 287 287  
FT CARBOHYD 415 415  
FT CARBOHYD 529 529  
FT CARBOHYD 532 532  
SQ SEQUENCE 586 AA; 67544 MW; FD88C62F5AF6C868 CRC64;  
  
Query Match 11.0%; Score 375; DB 1; Length 586;  
Best Local Similarity 23.1%; Pred. No. 3.6e-21;





OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;  
OC Apidae; Apis.  
OX NCBI\_TaxID=7460;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 18-41; 116-137 AND 255-267.  
RC TISSUE=Hypopharyngeal gland;  
RX MEDLINE=36192057; PubMed=8619864;  
RA Ohashi K., Sawata M., Takeuchi H., Natori S., Kubo T.;  
RT "Molecular cloning of cDNA and analysis of expression of the gene for  
RT alpha-glucosidase from the hypopharyngeal gland of the honeybee apis  
RT mellifera L.";  
RL Biochem. Biophys. Res. Commun. 221:380-385(1996).  
RN [2]  
RP CHARACTERIZATION.  
RC TISSUE=Hypopharyngeal gland;  
RX MEDLINE=97037073; PubMed=8882720;  
RA Kubo T., Sasaki M., Nakamura J., Sasagawa H., Ohashi K., Takeuchi H.,  
RA Natori S.;  
RT "Change in the expression of hypopharyngeal-gland proteins of the  
RT worker honeybees (Apis mellifera L.) with age and/or role.";  
RL J. Biochem. 119:291-295(1996).  
CC -1- FUNCTION: Converts sucrose in nectar to glucose and fructose.  
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing 1,4-  
CC linked D-glucose residues with release of D-glucose.  
CC -1- SUBUNIT: Monomer.  
CC -1- TISSUE SPECIFICITY: Expressed specifically in the hypopharyngeal  
CC glands of the forager (worker) honeybee.  
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; D79208; BAA11466.1; -.  
DR PIR; JC4714; JC4714.  
DR HSSP; P21332; IUOK.  
DR InterPro; IPR006589; Alp\_amy1\_cat\_sub.  
DR InterPro; IPR006047; Alpha\_amy1\_cat.  
DR Pfam; PF00128; alpha-amy1ase; 1.  
DR SMART; SMO0642; Amy; 1.  
KW Hydrolyase; Glycosidase; Glycoprotein; Signal.  
KW SIGNAL  
FT CHAIN 1 567 ALPHA-GLUCOSIDASE.  
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 247 247 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 319 319 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 499 499 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 507 507 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 567 AA; 65565 MW; A14D88DD657C99C2 CRC64;  
-----  
Query Match 9.6%; Score 325.5; DB 1; Length 567;  
Best Local Similarity 21.6%; Pred. No. 2.2e-17;  
Matches 128; Conservative 86; Mismatches 198; Indels 181; Gaps 23;  
-----  
QY 118 GDLKGLKDKIPYFQELGLTYIHIMPLFKCEPKSGDGYAVSSYRDVNPALGTIGDIREVI 177  
DB 49 GDIEGIKEKLDHFLFMGVDMFWISPIY--PSPMVDGFDISNYTVDHPIFGTISLDNLV 106  
QY 178 AALHEAGISAVVDIFNHTSNEHEWAQRCACAGDPLFDNFYIIFP-----DRMPDQYDRT 232  
DB 107 SAAHEKGLKITLDFVPHNTSDQHEWFQLSLKNIEPNYNIWHPGKIVNGKRV----- 160  
QY 233 IREIFPDQHPGFGFSQLEDGRVWT-----TFNSFQMDLNYSNPWFVRAMAGEMFL 283  
DB 161 -----FTNMVGVFG---GSAWSWREERQAYYLHQFAPEQPDNLNYNPVVLDDMQNVLRFW 212

QY 284 ANLGVDILRMDAVAFIWKQMTSC-----NLPOAHALIRAFNA 322  
DB 213 LRRGFDGFRVDALPYICEDMRFLDEPLSGETNDPNKTEYTLKIYTHDIPETVNVVRKFRD 272  
QY 323 VMRIAAPAVPFKSEAVIHPDOVQY--IGDEQCQIGYNPLQOMALLWNTLATREVNLLHQA 380  
DB 273 VLDEFPPQPKHMLIEAVYTNLSMTMKYDYGAD-----FPENFAFIKN--VSRDSN-----S 320  
QY 381 LTYRNLPEHTAWNVVVRSHDDIGWTFADEDAAYLGISGYDHRQFLNRFVVRFDGSFAR 440  
DB 321 SDFKKLVDN--WNTYMPPSGIPNVWPGNHDLRL-----VSRF----- 356  
QY 441 GVPEQYNPSTGDCRVSGTAALVGLAODDPHAYDRILKLYSIALSTGGLPLIYGDEVG- 499  
DB 357 -----GEEKRMITTMSLLLPGVAVNYGDEIGM 385  
QY 500 -----TLND-----DWSQDSN---KSDDSRWAHRPRYNE 526  
DB 386 SDTYISWEDTQDPQCGAGKENYQTMSPDPAFTPQMDSDVSAGFSSSNWTWL---RVNE 442  
QY 527 -----ALYAQRNDPSTAAGIYQGLRHMIAVRSNPRFDGRLVTENTN--NKHIIIGYIR 579  
DB 443 NYKTVNLAEEKDKNS-----FENMFKEFASLKKSPYEKANL--NTRMLNDNVFAFSR 494  
QY 580 ----NNALLAFGNFSEYEQVTYTAHTLOAMPFKAHDLIG-----KTVSLNODL 623  
DB 495 ETEDNGSLYAILNFSNEQIVDLKAFNNVPKLNMFEYNNFNSDIKSISNNEQV 547  
-----  
RESULT 7  
MGTA\_THEME  
ID MGTA\_THEME STANDARD; PRT; 441 AA.  
AC P80099;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE 4-alpha-glucanotransferase (EC 2.4.1.25) (Amylomaltase)  
DE (Disproportionating enzyme) (D-enzyme).  
GN MGTA OR TM0364.  
OS Thermotoga maritima.  
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.  
OX NCBI\_TaxID=2336;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MSB8 / DSM 3109 / ATCC 43589;  
RX MEDLINE=9287316; PubMed=10360571;  
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,  
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,  
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,  
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,  
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,  
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;  
RT "Evidence for lateral gene transfer between Archaea and Bacteria from  
RT Nature 399:323-329(1999).  
RL Nature 399:323-329(1999).  
RN [2]  
RP SEQUENCE OF 1-23, AND CHARACTERIZATION.  
RC STRAIN=MSB8 / DSM 3109 / ATCC 43589;  
RX MEDLINE=92331687; PubMed=1628664;  
RA Liebl W., Feil R., Gabelberger J., Kellermann J., Schleifer K.-H.;  
RT "Purification and characterization of a novel thermostable 4-alpha-  
RT glucanotransferase of Thermotoga maritima cloned in Escherichia  
RT coli.";  
RL Eur. J. Biochem. 207:81-88(1992).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).  
RX MEDLINE=22135897; PubMed=12139940;  
RA Roujeinikova A., Raasch C., Sedelnikova S., Liebl W., Rice D.W.;  
RT "Crystal structure of Thermotoga maritima 4-alpha-glucanotransferase  
RT and its acarbose complex: implications for substrate specificity and  
RL J. Mol. Biol. 321:149-162(2002).



```

CC -1- CATALYTIC ACTIVITY: Transfers a segment of a (1,4)-alpha-D-glucan
CC to a new 4-position in an acceptor, which may be glucose or (1,4)-
CC alpha-D-glucan.
CC -1- COFACTOR: Binds 1 calcium ion per subunit.
CC -1- SUBUNIT: Monomer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001716; AAD35451.1; -.
CC PDB; 1LMH; 31-JUL-02.
CC TIGR; TM0364; -.
CC InterPro; IPR006589; Alp_ amyl_cat_sub.
CC InterPro; IPR006047; Alpha_ amyl_cat.
CC InterPro; IPR006046; Glyco_hydro_13.
CC Pfam; PF00128; alpha-amylase; 1.
CC PRINTS; PR00110; ALPHAAMYLASE.
CC SMART; SM00642; Aamy; 1.
CC Carbohydrate metabolism; Transferase; Glycosyltransferase;
CC Calcium-binding; 3D-structure; Complete proteome.
CC ACT_SITE 186 186
CC ACT_SITE 216 216
CC ACT_SITE 278 278
CC METAL 13 13 CALCIUM.
CC METAL 15 15 CALCIUM.
CC METAL 17 17 CALCIUM.
CC METAL 19 19 CALCIUM. (VIA CARBONYL OXYGEN).
CC METAL 21 21 CALCIUM.
CC CONFLICT 1 1 M -> A (IN REF. 2).
CC SEQUENCE 441 AA; 51843 MW; 9FB4C2ABC9D3DF3A CRC64;
SQ
Query Match 9.5%; Score 324.5; DB 1; Length 441;
Best Local Similarity 23.2%; Pred. No. 1.9e-17;
Matches 126; Conservative 81; Mismatches 186; Indels 151; Gaps 21;
QY 107 VGVGVYVDLF-----AGDLKGLKDKIPYQELGLTYLHMLPKCPKSGSDGYAVSS 159
DB 2 IGYQIVYRSFRDGNLDVGDFGRGLKNVAVSYLKELGIDFWLMPEVF--SSISFHGYDVID 58
QY 160 YRDVNPALGTIGDLREVIALHEAGISAVVDFTFNHTSNHEHWAQRCAAGDPLFDNFYIT 219
DB 59 FYSFKAEGYSEREFKEMIEAFHDSGIKVLDLPIHHTGLHTWFQALKGDPHYRD-YYV 117
QY 220 FPDRRMPDQYDRTLREIFPDQ--HPGGFSQLDEGRWWTTFNSFQWDLNYSNPWVFRAMA 277
DB 118 WANKETDLDLDER--REWDGEKIWHF-----LEDGRFYRGLGFPSPDLINDNPQVFDEKK 169
QY 278 GEMLFLANLGVLDILRMDAVAFI-----WKQMGTSCE--LPQAHALIRAFNAVNR 325
DB 170 RLVLHLLDMGVDGFRFDAAKHEMDTIEQNVRFWKYLFLDLKGIFLAIRIWAARMVDEHGR 229
QY 326 IAAPAVFEKSEAIWHPDQVQYIGQDECQIGYNPLQMALIMNTLATREVNLLHQALTYRH 385
DB 230 IFGYMLNFDTSHCI-----KEAVKMKENTRVLIESIERAVIGKD 267
QY 386 NLPEHTAWVNYVRSDDIGWTFRADEDAAYLGISGYDHRQFLNRFVNRFDGSPARGVPPQ 445
DB 268 YLP-----VNFTSNHD-----MSR--LASFEGGFSK----- 291
QY 446 YNPSTGDCRVSGTAALVGLAODDPHAVDRIKLYSIALSTGGPLIYLIGDEVG----- 499
DB 292 -----EKIKLSISILFTLPGLVPLVFGDELGMKGVVYQ 323
QY 500 -----TLNDDDDWSQDSNKSDDSRWAHRRPYNE-----ALYAQRNDPSTAAGQIYQGLRH 548
DB 324 KENTBEVLLDPFPMNESMCVEGQTFWKM-PAYNGPFGSGISVEYQKRDPDSI-----LSH 375

```

[illegible]

RESULT 8	
016G_BACTR	
ID 016G_BACTR	STANDARD;
	PRT; 562 AA

DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Oligo-1,6-glucosidase (EC 3.2.1.10) (Oligosaccharide alpha-1,6-glucosidase) (Sucrase-isomaltase) (Isomaltase) (Dextrin 6-alpha-D-glucanohydrolase).  
DE MALT.  
GN *Bacillus thermoglucosidasius* (*Geobacillus thermoglucosidasius*).  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Geobacillus*.  
OX NCBI\_TaxID=1426;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=KP1006;  
RX MEDLINE=92105078; PubMed=1761534;  
RA Watanabe K., Chishiro K., Kitamura K., Suzuki Y.;  
RT "Proline residues responsible for thermostability occur with high  
RT frequency in the loop regions of an extremely thermostable oligo-1,6-  
RT glucosidase from *Bacillus thermoglucosidasius* KP1006.";

RL J. Biol Chem. 266:24287-24294(1991).  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,6-alpha-D-glucosidic linkages  
 CC in some oligosaccharides produced from starch and glycogen by  
 CC alpha-amylase, and in isomaltose.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- MISCELLANEOUS: THIS ENZYME IS EXTREMELY THERMOSTABLE.  
 CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC -----
DR EMBL; D10487; BAA01368.1; -.
DR PIR; A41707; A41707.
DR HSSP; P21332; 1UOK.
DR InterPro; IPR006589; Alp_amy1_cat_sub.
DR InterPro; IPR006047; Alpha_amy1_cat.
DR Pfam; PF00128; alpha-amy1ase; 1.
DR SMART; SM00642; Amy; 1.
KW Hydrolase; Glycosidase.
FT ACT_SITE 199 199 BY SIMILARITY.
FT ACT_SITE 256 256 BY SIMILARITY.
FT ACT_SITE 330 330 BY SIMILARITY.
SQ SEQUENCE 562 AA; 66505 MW; 925EB5924726D42A CRC64;

```

Query Match	9.2%;	Score 314;	DB 1;	Length 562;
Best Local Similarity	22.6%;	Pred. No. 1.7e-16;		
Matches 138; Conservative	90;	Mismatches 185;	Indels 198;	Gaps 30;

Dy 118 GDLKGLKDIPYFOELGTVYLHMLPKCKPEKGSDGGYAASSRYRDVNPALGTIGDIREVI 177  
||::| : | : ||:: : ||::| : ||::| : ||::| : ||::| : ||:  
  
Db 28 GDIRGI IAKLDYLKELGVDVAWLSPVKYSP - NDDNGIDISDRDIMDEFGTMAWMKTML 85

Oy 178 AALHEAGISA VDFIENHTSNEHW--AQRCAAGDPLEDNFYIF---PDRRMPDQYRT 232  
:||::| : | : ||::| : ||::| : ||::| : ||::| : ||::| : ||:  
  
Db 86 EEMKRGIKLVM DLVVNHNDSDEHPFIESRSKSNDPNYRD--YYIWPGKNGEKPNNWE-- 141



```

RT genes of a strictly anaerobic thermophile, Dictyoglomus thermophilum,
RT and their nucleotide sequences with extremely low
RT guanine-plus-cytosine contents."';
RL Eur. J. Biochem. 176:243-253(1988).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: Binds 1 calcium ion per subunit (By similarity).
CC -1- PATHWAY: Polysaccharide degradation.
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- MISCELLANEOUS: THE PROFILE OF PRODUCTION OF REDUCING SUGAR FROM
CC SOLUBLE STARCH BY THE ACTION OF AMYC WAS INTERMEDIATE BETWEEN
CC THOSE OF AMYA & AMYB.
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
-----
CC This swiss-prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outpost at -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X15948; CA034072.1; -.
DR PIR; S01313; S01313.
DR HSSP; P21332; 1UOK.
DR InterPro; IPR006589; Alp_amy1_cat_sub.
DR InterPro; IPR006047; Alpha_amy1_cat.
DR InterPro; IPR006046; Glyco_hydro_13.
DR Pfam; PF00128; alpha-amylase; 1.
DR PRINTS; PR00110; ALPHAAMYLASE.
DR SMART; SM00642; Aamy; 1.
KW Carbohydrate metabolism; Hydrolyase; Glycosidase; Calcium-binding;
KW Multigene family.
FT ACT_SITE 213 213 BY SIMILARITY.
FT ACT_SITE 248 248 BY SIMILARITY.
FT ACT_SITE 314 314 BY SIMILARITY.
FT METAL 127 127 CALCIUM (BY SIMILARITY).
FT METAL 183 183 CALCIUM (BY SIMILARITY).
FT METAL 217 217 CALCIUM (VIA CARBONYL OXYGEN) (BY
FT SIMILARITY).
SQ SEQUENCE 498 AA; 59063 MW; C53B3DDAD22074F5 CRC64;
Query Match 8.9%; Score 304; DB 1; Length 498;
Best Local Similarity 22.4%; Pred. No. 8.5e-16;
Matches 123; Conservative 83; Mismatches 208; Indels 134; Gaps 22;
QY 118 GDKGLKD KIPYFOELG LTYLHMLPFKEPKSGSDGYAVSSYRDVN PALGTIGDLREVI 177
Db 54 GDNLGLIDKL D YFKNLTIT ALWLPIF--PSVS YHGXDVTDYDIHPGYGT MEDFE NLI 110
QY 178 AALHEAGISAVVD FTFNHSTNE HEWAQRCAA--GDPLFDNFYYI PPDRRMPDQY RTLR E 235
Db 111 RKAHEKNIK ILDIVNVHTTSRHPEVSSASSYN SPYRD--YYIMSTEK----- 157
QY 236 IFPDQH PGFGFSQL EDGRVWTTFNSF QMDLNYSNPWFVRAMAGENULPLANLGVDILRM DA 295
Db 158 --PEKSNLMWK KP TG-YYYALFWSE MPDLNFDN PKVREEVKTA KFWEIEKGVDGF RLDA 214
QY 296 VAFI-----WKMG TSCENLPQAHAL IR AFNAVMRIAPAVFFKSEAIVHPDOVV 345
Db 215 AKHIYDDDSKNI QWWKE-----FYSLKS IKEDVYLLVG EWNDNEYK IA 257
QY 346 QYTGDEDCQ IGYNPLQM ALMNLTATREVNL LHQALT YRHNL PEHTAWNVYVRSH DDIG W 405
Db 258 EYXKGLPS NFNF-PLSDK IMNSSSKSR LR-----NYR IS-----RLKR LFGE NNNT 302
QY 406 TPADEDAAVLGISGY DHROFLNR FEVNR FDGSFARGVPFQYN PSTGD CRVSGTAAALV GL 465
Db 303 DFA--DAIFL---RNHDQ VRVF TF-FGG S----- 325
QY 466 AQDDPHAV DRIKLYSI ALSGTGL PLYLGD EV---GTL ND-----DWSQ DSNK SD S 516

```

```

Db      326  -----IDKSIAGSIYLLTAGNTFIYYGEEIGMEGSKPDEYIREPFKWTDDMSKYQT 378
QY      517  RWAHRPRYN-----EALYAQRNDPSTAAGIYQGLRHMIAVRQSNPRFDGRLVTENTNN 571
Db      379  YWI-IPRYNLPNGIALDTEEKDPNS---IYNHYKKLLEIRVKCRALSNKIERIKTOD 433
QY      572  KHIIGY---IRNNALLAFGNFSEYPTVTAHTLQAMPFKADHLIGKTVSLNOD-LTLQF 627
Db      434  RSLIAYKLELEDEKIMVHNLRIENTFNFN---EIKEKDILYIRNAKTEKXKIILGP 489
QY      628  YQVMMLLEI 635
Db      490  YSTVIYKI 497

RESULT 11
O16G_BACCE STANDARD; PRT; 558 AA.
ID O16G_BACCE
AC P21332;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Oligo-1,6-glucosidase (EC 3.2.1.10) (Oligosaccharide alpha-1,6-
DE glucosidase) (Sucrase-isomaltase) (Isomaltase) (Dextrin 6-alpha-D-
DE glucanohydrolase).
DE MALL.
OS Bacillus cereus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1396;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-10.
RC STRAIN=ATCC 7064;
RX MEDLINE=91006152; PubMed=2120057;
RA Watanabe K., Kitamura K., Iha H., Suzuki Y.;
RT "Primary structure of the oligo-1,6-glucosidase of Bacillus cereus
RT ATCC7064 deduced from the nucleotide sequence of the cloned gene.";
RL Eur. J. Biochem. 192:609-620(1990).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RC STRAIN=ATCC 7064;
RX MEDLINE=93380886; PubMed=8370659;
RA Kizaki H., Hata Y., Watanabe K., Katsube Y., Suzuki Y.;
RT "Polypeptide folding of Bacillus cereus ATCC7064
RT oligo-1,6-glucosidase revealed by 3.0-A resolution X-ray analysis.";
RL J. Biochem. 113:646-649(1993).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RC STRAIN=ATCC 7064;
RX MEDLINE=97336211; PubMed=9193006;
RA Watanabe K., Hata Y., Kizaki H., Katsube Y., Suzuki Y.;
RT "The refined crystal structure of Bacillus cereus oligo-1,6-
RT glucosidase at 2.0-A resolution: structural characterization of
RT proline-substitution sites for protein thermostabilization.";
RL J. Mol. Biol. 269:142-153(1997).
RN [4]
RP CATALYTIC ACTIVITY: Hydrolysis of 1,6-alpha-D-glucosidic linkages
RP in some oligosaccharides produced from starch and glycogen by
RP alpha-amylase, and in isomaltose.
RN [5]
RP SUBCELLULAR LOCATION: Cytoplasmic.
RN [6]
RP SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
RN [7]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RP between the Swiss Institute of Bioinformatics and the EMBL outstation -
RP the European Bioinformatics Institute. There are no restrictions on its
RP use by non-profit institutions as long as its content is in no way
RP modified and this statement is not removed. Usage by and for commercial
RP entities requires a license agreement (See http://www.isb-sib.ch/announce/
RP or send an email to license@isb-sib.ch).
RN [8]
RP EMBL; X53507; CAA37583.1; -
RN PIR; S13579; S13579.
RN PDB; 1UOK; 16-FEB-99.
RN InterPro; IPR006589; Alp_amy1_cat_sub.
RN InterPro; IPR006047; Alpha_amy1_cat.

```





QY 600 TLQAMPEKAHDLIGKTVSLNODLTLPQYVM 631  
Db 527 EVELL-IHNYDVENGPI-----ENITLRLPYEAM 553

RESULT 12  
MAL1\_DROME STANDARD; PRT; 508 AA.  
ID MAL1\_DROME STANDARD; PRT; 508 AA.  
AC P07191;  
DT 01-APR-1988 (Rel. 07, Created)  
DT 01-APR-1988 (Rel. 07, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Possible maltase D precursor (EC 3.2.1.20) (Larval visceral protein D).  
GN LVPD.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83216126; PubMed=6854639;  
RA Snyder M.; Davidson N.;  
RT "Two gene families clustered in a small region of the Drosophila genome."  
RL J. Mol. Biol. 166:101-118 (1983).  
RN [2]  
RP POSSIBLE FUNCTION.  
RX MEDLINE=88289361; PubMed=3135536;  
RA Henikoff S.; Wallace J.C.;  
RT "Detection of protein similarities using nucleotide sequence databases."  
RL Nucleic Acids Res. 16:6191-6204 (1988).  
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing 1,4-linked D-glucose residues with release of D-glucose.  
CC -1- DEVELOPMENTAL STAGE: One of the proteins expressed by the 44D cuticle gene cluster. Expressed in first, second and early 3rd instar larvae and in adults, but not in embryos or pupae.  
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; V00204; CAA23492.1; -.  
DR PIR; S08597; S08597.  
DR HSSP; P21332; 1UOK.  
DR FlyBase; FBgn0002569; lvpd.  
DR InterPro; IPR006589; Alp\_amy1\_cat\_sub.  
DR InterPro; IPR006047; Alpha\_amy1\_cat.  
DR Pfam; PF00128; alpha-amy1ase; 1.  
DR SMART; SM00642; Amy; 1.  
KM Hydrolase; Glycosidase; Signal.  
FT SIGNAL 1 23 POTENTIAL.  
FT CHAIN 24 508 POSSIBLE MALTASE D.  
FT CARBOHYD 30 30 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 312 312 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 508 AA; 57941 MW; C08411DED5A4DCEA CRC64;

Query Match 8.6%; Score 294; DB 1; Length 508;  
Best local Similarity 22.2%; Pred. No. 5.2e-15;  
Matches 119; Conservative 83; Mismatches 205; Indels 128; Gaps 19;

QY 118 GDLKGLKDKIPYRQELGLTYLHMLPFKCEKSDGAYAVSSYRDVNPALGTIGDLREVI 177  
Db 50 GDLKGLTIRLGLYKEIGITATWLSPIFTSP--MSDFGYDISNFYDIDPIFGTLEDFDDLI 107

QY 178 AALHEAGISAVVDIFNHTSNEHEMAQRCAAGDPLFDNFYIIPDRRMPDQYDRLREIF 237  
Db 108 VEAKSLGVKILDLFVFNHSSDENWFEKSVNREDDYDF-YWMDGKLNKEETGAR----- 161

QY 238 PDQHPGGSQLEDGRWVT-----TFNSFQMDLNTYSPWVFRAMAGEMFLANLGV 288  
Db 162 -DPPSNWVSFSGPMWTNNEKQOYFLHQFQVKQPDINTNPMVREHMLDVLKFMIDRGV 220

QY 289 DILRMDAVAFIWKQMGTCSCENLPQAHALIRAFNAVRIAPAVFEKSEAVIHPDQVQYI 348  
Db 221 DQFRIDAVPHIYEHRNADGS-----YFDEPVSGW 249

QY 349 GQDECQIGYNPLQWALLMNTLATRE---VNLHQALTYRHNLPHTAWVNYRSHDIG 404  
Db 250 GSDPNAYDYH-----DHITYKDQPATVDLMYEWREFLDNYRAQNGDSRVLLAE--A 299

QY 405 WTEADEDAAYLIGSYDHRQFLNRFVNRFDG-SFARGV--PFQYNPSTGDCRVSGTAA 461  
Db 300 YSSVETLSAYFGNSTHQGTQLEPMNFQMLYLSGYSTAKVGSIDYWMNT-MWKEHOTANW 358

QY 462 LVGLAQDDPHAVDR-----IKLYSIALSTGGLPLYLGDEVGLTND----- 504  
Db 359 VVG-NHDTNRVADRMGAHKYDLLNIVNALPGASVITYGEEIGMSNVDECTGDSCEDRD 417

QY 505 -----DWSQDSN---KSDSRMAH-RPRYNALYAQRNDPSTAAGIYQGLRHMAVVRQ 554  
Db 418 GERTPMQWTAGKNADFSDESTWLLPLSPYQRYNVQTERGSRSSLNIFKGLQELKS--- 474

QY 555 SNPRFDGRLVTFNTNNKHIIIGYIRNNALLAF---GNFSFYQTVTAHTLQAMPF 606  
Db 475 -----SSAFLAFKEDGFSY--EAVTBOVLQIIRY 502

RESULT 13  
AMY\_BACME STANDARD; PRT; 520 AA.  
ID AMY\_BACME STANDARD; PRT; 520 AA.  
AC P20845;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Alpha-amy1ase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).  
OS Bacillus megaterium.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1404;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=88262518; PubMed=2455281;  
RA Metz R.J.; Allen L.N.; Cao T.M.; Zeman N.W.;  
RT "Nucleotide sequence of an amy1ase gene from Bacillus megaterium."  
RL Nucleic Acids Res. 16:5203-5203 (1988).  
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic linkages in oligosaccharides and polysaccharides.  
CC -1- COFACTOR: Binds 1 calcium ion per subunit (By similarity).  
CC -1- SUBUNIT: Monomer (By similarity).  
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X07261; CAA30247.1; -.  
DR PIR; S01031; S01031.  
DR HSSP; P21332; 1UOK.  
DR InterPro; IPR006589; Alp\_amy1\_cat\_sub.  
DR InterPro; IPR006047; Alpha\_amy1\_cat.  
DR InterPro; IPR006046; Glyco\_hydro\_13.  
DR Pfam; PF00128; alpha-amy1ase; 1.

DR PRINTS; PR00110; ALPHAAMYLASE.  
DR SMART; SM00642; Amy; 1.  
KW Carbohydrate metabolism; Hydrolase; Glycosidase; Calcium-binding;  
KW Signal.  
FT SIGNAL 1 27  
FT CHAIN 28 520 ALPHA-AMYLASE.  
FT ACT\_SITE 233 233 BY SIMILARITY.  
FT ACT\_SITE 273 273 BY SIMILARITY.  
FT ACT\_SITE 340 340 BY SIMILARITY.  
FT METAL 140 140 CALCIUM (BY SIMILARITY).  
FT METAL 203 203 CALCIUM (BY SIMILARITY).  
FT METAL 237 237 CALCIUM (VIA CARBONYL OXYGEN) (BY  
SIMILARITY).  
SQ SEQUENCE 520 AA; 58761 MW; 2C649D957F5835C0 CRC64;  
Query Match 8.6%; Score 292; DB 1; Length 520;  
Best Local Similarity 22.8%; Pred. No. 7.6e-15;  
Matches 120; Conservative 74; Mismatches 208; Indels 124; Gaps 18;  
QY 112 YVDLF-----AGDLKGLKDKIPYF-----QELGLTYLHMLPFCPEKSGSDGYA 156  
DB 46 YVNSFYDANKDGHGDLKGLTQKLDYNDNSHTKNDLQVNGIWMWVFNPS---SYHKYD 102  
QY 157 VSSYRDVNPALGTIGDLREVIALHREAGISAVVDFFINHTSNEHEWAQRCAGDPLFDNF 216  
DB 103 VTDYNNIDPQYGNLQDFRKLMEKADKDYKIMDLVNNHTSSEHPWFOALDKNSKYRD 162  
QY 217 YYIFPRRMPDQYDRTLREIFPDQHGGSQ----LEDGRWVTTTNSFQWDLNYSNPW 271  
DB 163 YYIWDK-----NTDLNE-----KSGWQQVWHKAPNGEYFYGTFWEGMPDLNYPN 210  
QY 272 VFRAMAGMLFLANLGVDLIRMDAVAFIWKMGTSCENLPOAHALIRAFNAVRIAPAV 331  
DB 211 VRKEMINVGKFWLKQGVDFRLDALHIFK--GQTPGAKKNILWNEFRDAMKENPNV 268  
QY 332 FFKSEALVHPDQVQYIGQDECQIGYNPLQWALLMNTLATREVNLLHQALTYRHNLPEHT 391  
DB 269 YLTGEVMDQPEVVAFY-----YQSLDSLNFNFLAG 298  
QY 392 AMVNVYRSHDDIGMTFADEDAAYLGISGYDHRQFLNRFVNRFDGSEFARGVPFQYNPSTG 451  
DB 299 KIVSSVYKAGNDG-----IATAAATDELFKSYNPNKIDGIFLT---NHDQNRV 344  
QY 452 DCRVSGTAALVGLAODDPHAVDRIKLYSIALSTGGLPLIYLGDEVGTLD--DD----- 505  
DB 345 MSELSD-----VNKAASAAILLTLPNPIYIGEIGMTGEKPELDLIRE 390  
QY 506 ---WSQDSNKSDDSRWAHRRPN-----EALYAQRNDPSTAGQIYQGLRHMIARQ 554  
DB 391 PFRW-YEENGIGQTSW-ETPVYNGKGVSEVAQTOKD-----SLNHYREMIRVRQ 441  
QY 555 SNPRFDGRLVTENTNKKHIIIGYIR---NNALLAFGNFSEYPTVT 597  
DB 442 QHEELVKGTLQGISVDSKEVAVSRTYKGSISVYHNISNQPVKVS 487  
RESULT 14  
016G BACCO STANDARD; PRT; 555 AA.  
AC Q45101;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Oligo-1,6-glucosidase (EC 3.2.1.10) (Oligosaccharide alpha-1,6-  
glucosidase) (Sucrease-isomaltase) (Isomaltase) (Dextrin 6-alpha-D-  
glucanohydrolase).  
GN MALL.  
OS Bacillus coagulans.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1398;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 7050;

RX MEDLINE=96236197; PubMed=8787404;  
RA Watanabe K., Kitamura K., Suzuki Y.;  
RT "Analysis of the critical sites for protein thermostabilization by  
RT proline substitution in oligo-1,6-glucosidase from Bacillus coagulans  
RT ATCC 7050 and the evolutionary consideration of proline residues.";  
RL Appl. Environ. Microbiol. 62:2066-2073 (1996).  
CC -I- CATALYTIC ACTIVITY: Hydrolysis of 1,6-alpha-D-glucosidic linkages  
CC in some oligosaccharides produced from starch and glycogen by  
CC alpha-amylase, and in isomaltose.  
CC -I- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -I- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D78342; BAA11354.1; -.  
DR HSSP; P21332; IUOK.  
DR InterPro; IPR006589; Alp\_amy1\_cat\_sub.  
DR InterPro; IPR006047; Alpha\_amy1\_cat.  
DR Pfam; PF00128; alpha-amylase; 1.  
DR SMART; SM00642; Amy; 1.  
KW Hydrolase; Glycosidase.  
FT ACT\_SITE 199 199 BY SIMILARITY.  
FT ACT\_SITE 255 255 BY SIMILARITY.  
FT ACT\_SITE 332 332 BY SIMILARITY.  
SQ SEQUENCE 555 AA; 64990 MW; CE2B81D7B4E226D8 CRC64;  
Query Match 8.5%; Score 289.5; DB 1; Length 555;  
Best Local Similarity 21.7%; Pred. No. 1.3e-14;  
Matches 133; Conservative 79; Mismatches 220; Indels 181; Gaps 23;  
QY 118 GDLLKGLKDKIPYFQELGLTYLHMLPFCPEKSGSDGYAVSSYRDVNPALGTIGDLREVI 177  
DB 27 GDLRGIMDKLDYLTGLGIDCIWISPYDSPQ--DDNGYDIRDYRKIDKMGFTNEDMDRL 84  
QY 178 AALHEAGISAVVDFFINHTSNEHEW--AQRCAAGDPLFDNFYIFPDRRMPD----- 227  
DB 85 DEAHARGIKIVMDLVNHTSDEHAFVESRKSXDNPYRD--FYFWKDPK-PDGTPPNNWG 141  
QY 228 -----QYDRTLREIFPDQHGGSQLEDGRWVTTTNSFQWDLNYSNPVFRAMAGE 279  
DB 142 SMFSGSAGEYDET-----TGQYLLHYFSKKQPDILWNEKVRKEIYDM 184  
QY 280 MLFLANLGVDLIRMDAVAFIWK-----QMGTSCENLPOAHALIRAFNAV 324  
DB 185 MKFWMQGVYDGRMDVIGISKFLDPPEYELPEGQKYGIGKYHANQRLHAFTQEMNR-- 242  
QY 325 RIAAPAVFFKSEALVHPDQVQYIGQDECQIGYNPLQWALLMNTLATREVNLLHQALTYR 384  
DB 243 -----EVLISKYDCMT---VGELIGSD-----VEIARKYTGPDRLHLM 278  
QY 385 HNLPEHTAWNVYRSHDDIG-----WTFADEDAAYLGISGYDHRQFLNRFVNRFD 435  
DB 279 -----PNF--EHMDVDTPKPSGAKKALKPFDLVEL-----KQILSRWQYELAD 320  
QY 436 GSFARGVPQYNPSTGDCRVSGTAALVGLAODDPHAVDRIKLYSIALSTGGLPLIYLG 495  
DB 321 TGM-NALYFENHDQ-----ARVSRWGNDDTYRAECAKAPATILHGLKGTPTIYQG 370  
QY 496 DEVGTLDND-----DWS 507  
DB 371 EEIGWVADLELEYDDIEIRNAYQELVMENQIMSKDEFLTAVRKGRDNARTPMQWDGS 430  
QY 508 QDSNKSDDSRWAH-RPRYNALYAQRNDPSTAGQIYQGLRHMAVQSNPRFDGRLVT 566  
DB 431 FNAGFTGTPLKVNRSYSEINVAKALQEPDSIFYYQSL---IKLRHSYDVFTDGRYEL 487  
QY 567 FNTNNKHIIIGYIRNN---ALLAFGNFSEYPTVTAHTLQAMPFKADLLIGKTVSLNODL 623

Db 488 IMPDHPHLYVYTRNESEKLVANLISE--NTVSFD---QPDNWKLLGNYEDTGTST 541  
QY 624 TLQPYQ--VMLE 634  
Db 542 LFRPYEAAYYLE 554

RESULT 15  
DEXB STREQ

ID DEXB STREQ STANDARD; PRT; 537 AA.

AC Q59905;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Glucan 1,6-alpha-glucosidase (EC 3.2.1.70) (Dextran glucosidase)  
DE (Exo-1,6-alpha-glucosidase) (Glucodextranase).  
GN DEXB.  
OS Streptococcus equisimilis.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=119602;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H46A;  
RX MEDLINE=94049672; PubMed=8232196;  
RA Mechold U., Steiner K., Vettermann S., Malke H.;  
RT "Genetic organization of the streptokinase region of the  
RT Streptococcus equisimilis H46A chromosome.";  
RL Mol. Gen. Genet. 241:129-140(1993).

CC -1- FUNCTION: The physiological substrates may be short  
CC isomaltosaccharides (By similarity).

CC -1- CATALYTIC ACTIVITY: Hydrolysis of (1->6)-alpha-D-glucosidic  
CC linkages in 1->6-alpha-D-glucans and derived oligosaccharides.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; X72832; CAA51348.1; -.

DR PIR; S39970; S39970.

DR HSSP; P21332; IUOK.

DR InterPro; IPR006589; Alp\_amy1\_cat\_sub.

DR InterPro; IPR006047; Alpha\_amy1\_cat.

DR Pfam; PF00128; alpha-amy1ase; 1.

DR SMART; SM00642; Amy; 1.

KW Hydrolase; Glycosidase.

SO SEQUENCE 537 AA; 61733 MW; 154DF0ACAF302FC7 CRC64;

Query Match 8.5%; Score 289; DB 1; Length 537;  
Best Local Similarity 22.0%; Pred. No. 1.4e-14;  
Matches 127; Conservative 89; Mismatches 236; Indels 124; Gaps 23;

QY 118 GDLKGLKDKIPYFQELGLTYLHLMPLFKCPEGKSDGGYAVSSYRDVNPALGTIGDLREVI 177  
Db 28 GDLKGLKDKIPYFQELGLTYLHLMPLFKCPEGKSDGGYAVSSYRDVNPALGTIGDLREVI 177  
QY 178 AALHEAGISAVVDLFFNHTSNEHEWAQRCAGDPLFDNFYIIFPDRRMPDQYDRTLEIF 237  
Db 86 AALHEAGISAVVDLFFNHTSNEHEWAQRCAGDPLFDNFYIIFPDRRMPDQYDRTLEIF 237  
QY 238 PDQHPGFSQLED--GRWVWTFNSFQWDLNYSNPWVERAMAGEMLFLANLGVDTLRMDA 295  
Db 140 ----SGSAWELDEASGYLHLFSKKQPDLMNENAHVRQKIYDMNMFYAKIGGFRMDV 195  
QY 296 VAFIWKQMGTSCE-NLPQAHALIRAFNAVRIAPAVFEKSEAIVHPDQVQYIGQDECO 354  
Db 140 ----SGSAWELDEASGYLHLFSKKQPDLMNENAHVRQKIYDMNMFYAKIGGFRMDV 195

Db 196 IDLIGKIPDSEITGNGPRLHDYKEMN-----QATFGNHD-----VMTVGE---T 237  
QY 355 IGYNPLQMALMNTLATREYNIL---HQALTYRHLPEHTAWVNYVRSHDDIGWTEADE 410  
Db 238 WGAHP-EIARQYSRPENKELSMVFQFEHVGLQHKPNAPK-----WDYAE 281  
QY 411 ---DAYLIGISGYDHRQL---NRFVNRFD-----GS-----F 438  
Db 282 LDVPALKTIFSKWQTELKLGEGWNSLFWNNHDLPRVLSIWGNDISYREKSAKALALLHL 341  
QY 439 ARGVPFQYNPSTGDCRVSGTAAALVGLAODDPHAVDRIKL-YSIALSTGGLPLIYLGE 497  
Db 342 MRGTPYIY-----QGEIGMTNYPFKDLTEVDDIESLNTAKEMENGVPARVMSS 392  
QY 498 VGTINDDD-----WSQDS-----NKSDDSRWAHRPRYNEALYAQRNDPSTAAGQIYQGLR 547  
Db 393 IRKVGGRDNARTPMQWSKDTAGFSEAQETWLPVNPNYQELNVA---DALANQDSIFYTYQ 449  
QY 548 HMLAVRQSNPRFDGRLVTFTNTNNKHIIIGYIRNNALLAFGNSESEYQGTVAHTLQAMPFK 607  
Db 450 QLTALRKDQDWLVEADYHLLPPTADK-VFAYQR-----QFGEETVIVVNSDQEQVF 500  
QY 608 AHDLIGKTVSLNQDLT-----LQPYQVMLEIA 636  
Db 501 AKDLAAGAEVITNTDVKVLETKHLQPWDACVXLS 536

Search completed: May 25, 2004, 19:31:18  
Job time : 43 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 19:24:11 ; Search time 20 Seconds  
(without alignments)  
3058.891 Million cell updates/sec

Title: US-09-807-146-1

Perfect score: 3408  
Sequence: 1 MLPTQOVGLILQYKTRIL.....VSLNQDILQPYQVMWLELA 636

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

```
Database : PIR_78:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Query				DB	ID	Description
No.	Score	Match	Length				
1	1319.5	38.7	644	2	C75457		alpha-amylase - De
2	919.5	27.0	584	2	C87390		amylsucrase limpo
3	541.5	15.9	1100	2	G83376		probable trehalose
4	520.5	15.3	552	2	E75322		probable trehalose
5	509.5	15.0	601	2	G70983		probable glycoseyl
6	504	14.8	573	2	S71450		trehalose synthase
7	452.5	13.3	568	2	H64879		probable membrane
8	445	13.1	544	2	C95854		probable trehalose
9	393	11.5	566	2	AF1803		Sucrose phosphoryl
10	386.5	11.3	434	2	B85754		probable polysacch
11	373	10.9	560	2	AF1416		Sucrose phosphoryl
12	333.5	9.8	548	2	AF1231		alpha, alpha-phosph
13	330.5	9.7	549	2	S44188		alpha-glucosidase
14	325.5	9.6	441	2	S60618		4-alpha-glucanottra
15	325.5	9.6	567	2	JC4714		alpha-glucosidase
16	321.5	9.4	549	2	C89930		alpha-D-1,4-glucos
17	318.5	9.3	548	2	AE1585		alpha, alpha-phosph
18	314	9.2	562	2	A41707		oligo-1,6-glucosid
19	309.5	9.1	541	2	H98083		dextranase (EC 3.2
20	307	9.0	587	2	S42358		alpha-glucosidase
21	304	8.9	498	2	S01313		alpha-amylase (EC
22	304	8.9	558	2	S13579		oligo-1,6-glucosid
23	302.5	8.9	564	2	H75403		glycosyl hydrolase
24	297.5	8.7	541	2	B95220		dextran glucosidas
25	295	8.7	546	2	F70866		probable maltase -
26	294	8.6	508	2	S08597		hypothetical prote
27	293.5	8.6	599	2	S55363		maltase-like prote
28	292	8.6	520	2	S01031		alpha-amylase (EC
29	289	8.5	537	2	S39970		glucan 1,6-alpha-g

30	288.5	8.5	562	2	C82263	trehalose-6-phosph
31	287.5	8.4	498	2	S55362	maltase-like prote
32	286.5	8.4	561	2	AE2649	alpha-glucosidase
33	286.5	8.4	594	2	D97431	probable alpha-glu
34	286	8.4	555	2	AH0449	alpha, alpha-phosph
35	285	8.4	522	2	S07253	hypothetical prote
36	285	8.4	534	2	T30268	oligo-1,6-glucosid
37	281	8.2	505	2	S08598	hypothetical prote
38	281	8.2	584	2	S46183	alpha-glucosidase
39	280.5	8.2	551	2	H91280	trehalase 6-P hydr
40	280.5	8.2	551	2	H86121	trehalase 6-P hydr
41	280	8.2	584	1	ALBY	alpha-glucosidase
42	280	8.2	584	2	S64627	alpha-glucosidase
43	279.5	8.2	550	2	AE1058	alpha, alpha-phosph
44	278	8.2	553	2	AI1097	oligo-1,6-glucosid
45	277	8.1	553	2	AH1460	oligo-1,6-glucosid

## ALIGNMENTS

```

RESULT 1
C75457
alpha-amylase - Deinococcus radiodurans (strain R1)
C/Species: Deinococcus radiodurans
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C/Accession: C75457
R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
, M.; Shen, M.; Yamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
Science 286, 1571-1577, 1999
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A/Reference number: A75250; MUID:20036896; PMID:10567266
A/Accession: C75457
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-644 <WHI>
A/Cross-references: GB:AE001946; GB:AE000513; NID:g6458655; PIDN:AAF10510.1; PID:g6458655
A/Experimental source: strain R1
C/Genetics:
A/Gene: DR0933
A/Map position: 1
C/Superfamily: alpha-glucosidase; alpha-amylase core homology

Query Match 38.7%; Score 1319.5; DB 2; Length 644;
Best Local Similarity 42.1%; Pred. No. 8e-98;
Matches 281; Conservative 110; Mismatches 215; Indels 61; Gaps 14;

QY 1 MLPTQOVGLILQYLKRLDIYTPQAGIEKSEDMRQFSRRMDTHFPKLNELDSVYG 60
Db 1 MLTPD-----LAARL-----RLAFDDRDRAETRLRLERYGPFLADNLRAVYG 43

QY 61 NN-EALL-PMLBMLLAQAWQSYSQRNSSLKQIDILARENPNPWILSNKQVGVCYVDLFAG 118
Db 44 NHADALLIGELLEVM-----HAYHARPADLKRLEDEARLLRPDWLQGPBMWGYAVYDRFAG 99

QY 119 DLKGLKDKIPYFQELGLTYLHMLPLFKCPEKSDGGAIVSSYRDVNPALGTTGDLREVI 178
Db 100 TLRGVGERLEYLGLGVTYLHLPLLRPRDGENGGYAVQDYRSVRPDLGTTIDDLISALAR 159

QY 179 ALHEAGISAVVDLFINHTSNEHEWAQRCACAGDPLFDNFYIIFPDRRMPDQYDRTLREIFP 238
Db 160 ELRGRLSLVLDLVNHFVAEEHEWAVRATAGEAAYRDYFHIFPDRTPQDAYERTLPEIFP 219

QY 239 DQHPGGFS-QLIEDGRWTTTFNSFQWDLNYSNFWFRAMAGEMLFLANLGVDLRMDAVA 297
Db 220 DPAFGNFTWNGBAGGWTTTFNRSQWMDVNWGNPAVFRLEYLDLILTLANRGVEYERLDAIA 279

QY 298 FIWKQMGTSCEMLPQAHALIRAFNAVMRILAAPAVFEKSEALIVHPDQVQYIGQDE----- 352
Db 280 FLMKRLGTDCQNGPEVHRLTHALRAATRIYAPAVAFAKAEAIYAPGDLIHLYGSRDHGRV 339

QY 353 CQIGYNPLQMLMLNLTATREVNLIHQALTYRHNLPEHTAWNVYVRSHDIGMTFADEDA 412

```



Db 340 SDMAHNSIMVQLWSSSLASRDTRLTLAALAAFPKPTNTTWGYVCHDDIGWAIADEDA 399  
Qy 413 AYLGISGYDHRQFLNRFVNRFDGSPARGVFPQYNPSTGDCRVSGTAALVGL----- 465  
Db 400 ARVGLSGPARHRLSDFSYSGEFGSFAAGLVFQHHPQTGDRIISGTAASLAGLDLLETG 459  
Qy 466 -AODDPAHVDRIKLYSIALSTGGLPLYLGDEVGTLNDDDWSQDSNKSDDSRWAHRPRY 524  
Db 460 DAERVNDALARLLHLHAVMLGFGVPLLYMGDELALNDTFAAVPAHAADNRWVHRPQM 519  
Qy 525 NEALYAQ-----RNDPSTAAGQIYQGLRHMIAVRQSNPRFDGGRLVTFNTNKH----- 574  
Db 520 DWELVASQAADAATGQPTVPAGRMFAGLRHLLAVRRTPHLHA-----STESRPLPSPD 573  
Qy 575 --IGYIR---NNALLAFGNFSEYPTVTAHTLQ-AMPEKAHDLIGKTVSL-NODLTLQ 626  
Db 574 PCVLLRREHPTGVLLQVYNFSEHHITPPTWPLQEQLGAVAHDLIGESQFHLGGPDIALE 633  
Qy 627 PYQVMWL 633  
Db 634 PYRALMWL 640

RESULT 2  
C87390  
amylosucrase [imported] - Caulobacter crescentus  
C/Species: Caulobacter crescentus  
C/Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C/Accession: C87390  
R;Niernan, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eise, J.; Heidelberg, J.  
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A/Title: Complete Genome Sequence of Caulobacter crescentus.  
A/Reference number: A87249; MUID:21173698; PMID:11259647  
A/Accession: C87390  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-584 <STO>  
A/Cross-references: GB:AE005673; NID:q13422449; PIDN:AAK23119.1; GSPDB:GN00148  
C/Genetics:  
A/Gene: CC1135

Query Match 27.0%; Score 919.5; DB 2; Length 584;  
Best Local Similarity 35.2%; Pred. No. 1e-65;  
Matches 215; Conservative 100; Mismatches 246; Indels 49; Gaps 16;  
Qy 44 MDTHFPKIMNELDSVYGNNEALPMLMLLAQAWQSYSQRNSSIKDIDDIARENPDWILS 103  
Db 1 MESRFAPK-----YGADARGPAVERLRLKTNLLKAAHARPEPLRALDARAADPAWLAHA 53  
Qy 104 NKQVGVGVYVDLFAAGDLKGLKDKIPYQELGLTYLHMLPLFKCEKSGDGYAVSSYRDV 163  
Db 54 PGQTATFYVDRAFAGDLNGVRGKLDYLTGLVRLHPLPLEPRPGSDSGFAVADYRKV 113  
Qy 164 NPALGTIGDLREVIALHEAGISAVVDFIFNHTSNEHEWAQRCAGADPLFDNFYIFPDR 223  
Db 114 DPLRGITDDLEALAGDLRQDRMDGLIDVVCNHTAREHAWAAKARAGDPAYRDYIYLPDA 173  
Qy 224 RMPDQYRTRLREIFPDQHPGFS-QLEDGRVWTTTENSFQWDLNYSNPWVERAMAGMLF 282  
Db 174 QSAARDBRELIDVFPDTPAGSFTYDAMGVWTTFFYPQWDLNYPANPAVFAEMLEVLIF 233  
Qy 283 LANLGVDIRMDAVAFIWKQMGTSCEMLPOAHALIRAFNAVMRIAPAVFKESEATVHPD 342  
Db 234 LAAKGAQGFRLDSAPFLWKQAGTTCRNIPQTYEIVEAWRAALSTIVAPGVLLAAEAIESV 293  
Qy 343 QVVOYITGDE--CQIGYNPLQMALMNTLATREVNLLHQALTYRHNLPEHTAMVNYVRSH 400  
Db 294 DVLPEFGESGNCLAYNNVMTALMAALADGDAVIARRCLAVARAKPAQAGAMLNYYRCH 353  
Qy 401 DDIGWTFADEDAAYLGISGYDHRQFLNRFVNRFDGSPARGVFPQYN-----PSTGDCRV 456

Db 354 DDLIWNAL--AAAYAPAS--DLRWSNAY--GNNEG-FSRGRAFOTAEGVPEST-----N 400  
Qy 457 GTAALVGLAQD--DPAVDRIKLYSIALSTGGLPLYLGDEVGTLNDDDWSQDSNKS 513  
Db 401 GMAAALAGLTADADGDCIGARRLRLLYGI IHALDGPPLYMGDEIGLDNDEAYQDDPLRA 460  
Qy 514 DDSRWHRPRYNEALYAQRNDPSTAAGQIY-----QGLRHMIAVRQSNPRFDGGRLV 566  
Db 461 GDGRWLRHPQMDWSLAERRERAGALQADLFATFARLGQRARRLATLGVAGP---ARPV- 515  
Qy 567 FNTNNKHIIIGYIRNNA---LLAFGNFSEYPTVTAHTLQAMPEKAHDLIGKTVSLNODL 623  
Db 516 -EYSSPAVLAFLRDEGARPFCLVANVSADAPQDF--ELPPAFAQABDVLDG-APSPAGAV 571  
Qy 624 TIQPYQVMWL 633  
Db 572 SLPPYGITWL 581

RESULT 3  
G83376  
probable trehalose synthase PA2152 [imported] - Pseudomonas aeruginosa (strain PA01)  
C/Species: Pseudomonas aeruginosa  
C/Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C/Accession: G83376  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidig, K.; Lim  
. J.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path  
A/Reference number: A82950; MUID:20437337; PMID:10984043  
A/Accession: G83376  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-1100 <STO>  
A/Cross-references: GB:AE004642; GB:AE004091; NID:g9948163; PIDN:AAG05540.1; GSPDB:GN00  
A/Experimental source: strain PA01  
C/Genetics:  
A/Gene: PA2152

Query Match 15.9%; Score 541.5; DB 2; Length 1100;  
Best Local Similarity 26.9%; Pred. No. 6.3e-35;  
Matches 165; Conservative 90; Mismatches 223; Indels 135; Gaps 25;  
Qy 96 NNPDWILSNKQVGVY---VDLF-----AGDLKGLKDKIPYQELGLTYLHMLPLFK 145  
Db 11 NDPQWYKD-----AVIYQVHYKSFYDANNNGIGDFAGLIEKLDYIADLGVTNLWLPLFY- 64  
Qy 146 CPEKSGDGYAVSSYRDVNPALGTIGDLREVIALHEAGISAVVDFIFNHTSNEHEW--- 202  
Db 65 -PSPRRDDGYDIAQYRGVHSDYGLADARRFIAEAHRRLRVITELVINHTSDQHPWFIR 123  
Qy 203 AQRCAAGDPLFDNFYIFPDRRMPDQYDRTLREIFPDQHPGFSQLEDGRVWWT----- 256  
Db 124 ARHAKKGSRRAD--YYWSD--SDEKYQGT-RIFID-----TEQSNWTWDPVAQY 170  
Qy 257 ---TFNSFQWDLNYSNPWVERAMAGEMFLANLGVDIRMDAVAFIWKQMGTSCEMLPOA 313  
Db 171 YWHRFYSHQPDLPDNPQVLRREVLGVMRYWLDMGVDGLRLDAIPYLIERDGTSSENLPET 230  
Qy 314 HALIRAFNAVMRIAPAVFKESEATVHPDQVVOYIG-----QDECQIGYNPLQMALM 367  
Db 231 HQVLRIRAEIDAHYPDRMLAEANQWPEPTRPYFGGEDGEGDECHMAFHPPLMPRM 290  
Qy 368 TLAATREVNLLHQALTYRHNLPEHTAMVNYVRSHDDIGWTFADEDAAYLGISGYDHRQFLN 427  
Db 291 AIAQEDRYEPTIDLRQTPDIPANCQWAIFLRNHDELTEMYTD-----DERDYL- 339  
Qy 428 RFEVNRFDGSPARGVFPQYNPSTGDCRV--GTAALVGLAQDDPFAVDRIKLYSIAL 484  
Db 340 -----WNHYAADRARLNLGIRRLAPIVERDR--RIELHSLTL 378

[illegible]

## RESULT 4

probable trehalose synthase - *Deinococcus radiodurans* (strain R1)  
C/Species: *Deinococcus radiodurans*  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
C/Accession: E75322  
R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
Science 286, 1571-1577, 1999  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
A/Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.  
A/Reference number: A75250; MUID:20036896; PMID:10567266  
A/Accession: E75322  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-552 <WHI>  
A/Cross-references: GB:AE002040; GB:AE000513; NID:g6459824; PIDN:AAF11586.1; PID:g6459822  
A/Experimental source: strain R1  
C/Genetics:  
A/Gene: DR2036  
A/Map position: 1  
C/Superfamily: alpha-glucosidase; alpha-amylase core homology

### Query Match

Query Match	15.3%;	Score 520.5;	DB 2;	Length 552;
Best Local Similarity	26.4%;	Pred. No. 1.1e-33;		
Matches 151; Conservative	91;	Mismatches 226;	Indels 103;	Gaps 18;

[illegible]

Db 413 GGFSTAQPSDCEFFPPIQDPVYGFGRVNVQSOLQDPSS----LLKWTARQLELRRAHPAFA 468

QY 561 GGRLVTENTNNKHIIIGYIR--NNALLAFGNFSEYPPQVTVAHTIQAMPFKAHDLIGKTV 617

Db 469 HGDLTFIETGNPAIIAFTROYDGETLLIVSNFAGNAQ---AGLLDLAPF-----VGRAPV 520

QY 618 SLN-----QDLTLQPYQVMWLEI 635

Db 521 TLGGASPLPVVTGNGQYFPVVMGKYDYMWLR 551

## RESULT 5

probable glycosyl hydrolase - Mycobacterium tuberculosis (strain H37Rv)  
C;Species: Mycobacterium tuberculosis  
C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C;Accession: G70983  
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998  
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A;Reference number: A70500; MUID:98295987; PMID:9634230  
A;Accession: G70983  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-601 <COL>  
A;Cross-references: GB:Z96071; GB:AL123456; NID:g3242254; PIDN:CAB09454.1; PID:g2181968  
A;Experimental source: strain H37Rv  
C;Genetics:  
A;Gene: RV0126  
C;Superfamily: alpha-glucosidase; alpha-amylase core homology

Query Match

Query Match	15.0%;	Score 509.5;	DB 2;	Length 601;
Best Local Similarity	27.6%;	Pred. No. 9.7e-33;		
Matches 160;	Conservative 83;	Mismatches 216;	Indels 121;	Gaps 24;

[illegible]

Db 488 MLAVRRRHPAFVAGAFQELGGSNPSVLAYVQVAGDDGDTVLGVNNLSRFPQPIELDLQ 547

QY 603 AMPEKAHDLIG-----GKTVSLNQDLTIQPYQVMLEI 635

Db 548 WTNYPVELLTGHVEFPRIQVPPYL--LTLPGHGFYWFQL 584

## RESULT 6

S71450  
trehalose synthase, maltose-specific - *Pimelobacter* sp. (strain R48)  
C/Species: *Pimelobacter* sp.  
A/Variety: strain R48  
C/Date: 09-Dec-1997 #sequence\_revision 09-Dec-1997 #text\_change 20-Jun-2000  
C/Accession: S71450; S78124  
R/Tsunaki, K.; Nishimoto, T.; Nakada, T.; Kubota, M.; Chaen, H.; Sugimoto, T.; Kurimoto, Biochim. Biophys. Acta 1290, 1-3, 1996  
A/Title: Cloning and sequencing of trehalose synthase gene from *Pimelobacter* sp. R48.  
A/Reference number: S71450; MUID:96239784; PMID:8645698  
A/Accession: S71450  
A/Molecule type: DNA  
A/Residues: 1-573 <TSU>  
A/Cross-references: EMBL:D78198, NID:g1536813, PIDN:BA11303.1, PID:g1536814  
A/Experimental source: strain R48  
A/Accession: S78124  
A/Molecule type: protein  
A/Residues: 6-25;47-63;188-198;235-241;354-373;526-545 <TSA>  
C/Genetics:  
A/Start codon: GTG  
C/Superfamily: alpha-glucosidase; alpha-amylase core homology  
F/6-573/Product: trehalose synthase #status experimental <MAT>

Query Match 14.8%; Score 504; DB 2; Length 573;

Best Local Similarity 25.6%; Pred. No. 2.5e-32;  
Matches 155; Conservative 96; Mismatches 232; Indels 122; Gaps 23;

```
QY PDWI-----LSNKÖVGVCVYDULFAGDLKGLKDKIPIFYÖELGTYLHLMPLFK 145
Db PEMFRTAVEYEVLVRSEFRDENAG-----TGDFRGLAEKLIDYLÖWLGVDCLMVPPFFS 65
QY CPEGKSDBGYAVSSYRDVNPAFGTIGIREVIAALHEAGISAVVDFTENHTSNEHEW--A 203
Db SP--LRDGGYDVADYTGITLPEIGIVEDFHAFLDGAHERGIRVIIDFWNHTSDAHWPFOA 123
QY ÖRCAGADELFDNFYIIFPDRMPDÖYDTLREIFFPDÖHPGGFS-ÖLEDGRWWTTFNSFÖ 262
Db SRSDPDGPYG-D-FYWSD---TDELYÖDARVIFVDTEPSNWITWDÖTRGÖYYWHRFHHÖ 178
QY WDLNYSNPWFERAMAGEMLFLANLGVILRMDAVAFIWKÖMGTSCENLPOÄHALIRAFNA 322
Db PDLNFDPKVÖDAMLÄMAFAWLDMGLDGFRLDAVPYLYERPGTNGENLPETHEMLKRVRR 238
QY VMRIAPAVEFEKSEATVHPDÖVVÖYIGÖDEÖIG-----YNPLOMALLMNTLAT 371
Db FVDDNYFPDRVLLYEANÖMPTDVEVEYFGPERREDGTVVGPESHMAFHFPVMPRIF--MAV 295
QY REYNI--LHÖALTTRYHNLPEHTAWNVNYVSHDDIGWTF-ADEDAAYLGISGYDHROFLNR 428
Db RESREFPISIEIMEÖTPAIRPEGCÖWGIFLNHDELTLBMVTDEDRDDYM----- 343
QY FFVNRFDGSFARGVPFÖYNPSTGDCRVSGSTAALVGLAÖDDPHAVDRIKLYSIALSTGS 488
Db GEYAKDPRMKAN-----IGTRRLAPLLDND--TNÖIELFTALLSLPG 385
QY LPLIYLGEDEVGTLNDDWSÖDSN-----KSDDSRWA-----HRPRYNEALY--- 529
Db SPLVLYGDEIG-MGDNIWLGDRDGVRTPMÖRTPDRNVGFSAAATPKLHLPTIODPVYGÖ 444
QY AQRNDPSTAAGÖIYÖGLRHMIAVRÖSNRPFDGRLVTFNTNNKHIIIGYIR----- 579
Db SVNVEAOLENPSS----LLHWTIRMITHIRÖQDAFGLGTFFEDLGGSNPAVALSYRELPGD 500
QY --NNALLAFGNFSEYPÖTVTAHTILOAMPFXAH---DLIGG---KTVSLNQDLTLÖPYÖV 630
```

Db 501 GGDVILCVNNLSRFPQV--ELDIRKYEGRVPVELIGVFPFPAVGELPYLLTISGHF 557

QY 631 MWLEI 635

Db 558 YWPRLL 562

## RESULT 7

H64879  
probable membrane protein b1309 - Escherichia coli (strain K-12)  
C;Species: Escherichia coli  
C;Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002  
C;Accession: H64879  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of Escherichia coli K-12.  
A;Reference number: A64720; MUID:97426617; PMID:9278503  
A;Accession: H64879  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-568 <BLAT>  
A;Cross-references: GB:AE000229; GB:U00096; NID:g1787566; PIDN:AAC74391.1; PID:g1787568  
A;Experimental source: strain K-12, substrain MGI655  
C;Superfamily: alpha-glucosidase; alpha-amylase core homology  
C;Keywords: transmembrane protein  
F;216-232/Domain: transmembrane #status predicted <TM01>  
F;414-430/Domain: transmembrane #status predicted <TM02>

Query Match 13.3%; Score 452.5; DB 2; Length 568;

Best Local Similarity 26.7%; Pred. No. 3.4e-28;  
Matches 165; Conservative 98; Mismatches 260;  
Indels 95; Gaps 26;

[illegible]



```

QY      617  VSLNQD LTLQPYQVMLE 634
      : | ::||:||||::
Db      552  I-LGNEITLRPWQVMWK 568

```

RESULT 8  
C95854

probable trehalose synthase protein [imported] - Sinorhizobium meliloti (strain 1021) ma  
C/Species: Sinorhizobium meliloti  
C/Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
C/Accession: C95854  
R/Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A/Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo  
A/Reference number: A95842; MUID:21396508; PMID:11481431  
A/Accession: C95854  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-544 <KUR>  
A/Cross-references: GB:AL591985; PIDN:CAC48499\_1; PID:g1513997L; GSPDB:GN03167  
A/Experimental source: strain 1021, megaplasmid pSymB  
R/Galbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
A/Title: The composite genome of the legume symbiont sinorhizobium meliloti.  
A/Reference number: A96039; MUID:21368234; PMID:11474104  
A/Contents: annotation  
C/Genetics:  
A/Gene: Smb20099  
A/Genome: plasmid  
C/Superfamily: alpha-glucosidase; alpha-amylase core homology

Query Match 13.1%; Score 445; DB 2; Length 544;  
Best Local Similarity 25.4%; Pred. No. 1.3e-27;  
Matches 157; Conservative 90; Mismatches 216; Indels 154; Gaps 23;

```

QY      84 NSSLKIDIDAR--ENNPDWILSNKQVGVCYVDLFAAGDLKGLKDKIPYFQELGLTYLHM 141
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      9  SSVIYCIDVRRFSDGNGDGI-----GDFIGLKERVSYLHYLGIDCWLS 52

```

```

QY      142 PLEKPEGKSDGGYAVSSYRDVNPALGTIGDLREVIATLHEAGISAVWDFEINHTSNEHE 201
      | : | | | | : : | | | | : : | | | | : | | | |
Db      53 PFRKSP--FADNGYDVSDYYSIDPVLGTLDLDFLHAAGEGQIRVWVTLVANHNTSSEHP 110

```

```

QY      202 WAQ-----RCAAGDPLFDFNYFIFFPDRMRMEDQYDRTLREIFPDQ---HPGGSQLEDGR. 252
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      111 WFOARRDARCRFRD-----YYWWSPPPP-----VAPDNKTAFFPGSSS----- 150

```

```

QY      253 WWT-----TFNSFQWDLNYSNPEWVERAMAGEML-----FLANIGVDILRMDDAVA 297
      |||      |  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      151 -VMTYDDLAAQAYYFHKFRHFQPDILNAP-----AVRDELLRVVDYWTLTGVDSGRVDAAP 205

```

```
QY      298 FIWKQMGTSCENLPQAHALIRAFANVRIAPAVFEKSEAIIVHPDQQVVQTIGQDECOIGY    357  
       | : | : | : | : | : | : | : | : | : | :  
Db     206 FVIGETGIHADPRDPQGFLREMRKLVDGRRODGILLGEADLAPEKLRSTYFGGKDILLF   265
```

QY 358 NPLQMALLMNTLATREVNLLHQALTYRHNLPEHTAMVNVYRSHDDIGWTFADEDAAYLGI 417  
| : | : || : : || : | : | : | : | :  
Db 266 NFVLCAAFASLAGOKADPIGALSINPEPPHGWANFLRLDELINDLPEDI----- 320

```

QY      418  SGYDHRQFLNRFVNRFDGSFARGVPEQYNPESTGDCRVSGTAALVGLAQDDEHAVDRIK  477
      : | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      321  ---RBETFAAFAPDEEMRIYGRGIRRLAP-----MLDGNRA-----RIE  357

```

```
QY      478 LLYSIALSTGGLPLIYLGDVEGTINDD-----DWSQDSNKSDDSRWAHRPRY 524  
        | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :  
Db      358 LAFSLILSSPGVPLILYGDEIG-LGEDLSRPGREPVRVPMQNNAGSNAGFST--AQRAKL 414
```

QY 525 NEA-----LYAQRNDSTAGQIYQGLRHMIAVRQSINPRFDGRLVTFENTNN 571  
: : | | | : : | | : : : : : : : : : : : :

Db 415 VÖPVTIDGPFSEFKRVNVEÅQREDP-----GSLINKVRAMITLARRRHELFNKGPFVMLHTGD 470

QY 572 KHITGYIRNNALLAFGNFSEYPOVTYHTLÖ-----AMPFKAHLLTGKTVSL 619

Db 471 PALFA-----LAYSDETFL-FVVLHNLSDGKRAEELPGAIDÄRLNDITLGEVEVL 521

QY 620 NQD---LTLÖPYÖVMM 633

Db 522 SGGRLAIGLGPFGYVWL 538

RESULT 9  
AF1803

Sucrose phosphorylase homolog lin2973 [imported] - *Listeria innocua* (strain Clip11262)  
C;Species: *Listeria innocua*  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 17-May-2002  
C;Accession: AF1803  
R;Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A;Authors: Krefelt, J.; Kuhn, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maitournam, A.; Nook, C.; Schlutener, T.; Simoes, N.; Tlherrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AF1803  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-566 <GLA>  
A;Cross-references: GB:AL592022; PIDN:CAC98198.1; PID:g16415513; GSPDB:GN00178  
A;Experimental source: strain Clip11262  
C;Genetics:  
A;Gene: lin2973  
C;Superfamily: alpha-glucosidase; alpha-amylase core homology

Query Match	11.5%;	Score 393;	DB 2;	Length 566;
Best Local Similarity	24.7%;	Pred. No. 2.1e-23;		
Matches 151;	Conservative 114;	Mismatches 261;	Indels 86;	Gaps 25;

```

Qy      51  LMNEIDSVYGN--EALLPMLTLLAQAWQSYSQRNSSLKDIDIARENNPWLISNKQVG 108
          |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db      4  LRKLSRLYSEDVESLATRIEARVQT---KQKLVKRD-----KW--DEKDIV 48

```

```

QY      109 GVCAYVDLEAGDLKGLKDKIPYFQELGITYL-----HLMPLFKCPREGKSDGYAVSSYR 161
          :| | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      49 LITYGDQFKEE---SKTILPTFKKAVDRILKSTFEVHFLPPYPY---SSDDGSVIDYK 102

```

```
QY      162 DVNDALGTIGDLREVIALTHEAGISAVVDEIFNHTSNHEWAQRCAAGDPLDNFYY-IF 220  
        ||| | :::: | :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:  
Db      103 AVNDELGDWEDIKEM----EQSARIMFDVFCNHMSAKSEWFKRRLAGDKKEQNFFVEMD 157
```

```

QY      221 PDRMRPDQYDRTLREIFPDQHPGGSQLEDGRWVWTTFNSFQWDLNYSNPWYFRAMAGEM 280
        ||      |      |      |      |      |      |      |      |      |      |
Db      158 PD---TDLSSVTRPRATPVLTPEQASGKEG-YIWTTFSEDQIDLNFACPETLYKMIDVL 213

```

QY 281 LPLANLGVLDLRMDAVAFIWKQMGSCENTLPQAHLIRAFNAVVRIRAPAEFFKSEA-IV 339  
:  
Db 214 MYLEEGAEYVRLDAVGFMWKPVTSSIHLDETHEIVKLFRDLVDMAAPGTLITTEINVP 273

QY 340 HPDQVVQY-IGDECCQIGYN-PLQMALIW-----NTLATREVNLIHQALTYRHNLPE 389  
| | : : | : | | : | : |  
Db 274 HYDNISYFNGEKEAHMVMYQFPLPLVLHAIHHGNAEFLSNMAKNL-----ELPEG 324

```

Qy      390 HTAWVNVYRSHDDIGWTFAD--EDAAVLGISGYDHRQFLNRFVFNRFDSFARGVPEQY 446
          : : : | | | : : | : : : : : : : : :
Db      325 KRTFNFPLASHDGIGLNPFVRGITPEAEILALVDLKEGALVSYKQNPDGTKS---PYEI 381

```

```

QY      447  NESTGDCRVSGTAAALVGLAQDDPEHAVDRIKLLYSIALSTGGLPLIYLGDVEVGLNDDW 506
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      382  NTIYMD-----ALSKQADTDILRLSRFLVAHAVALMSTPGVPAVYQSTILGSRNDYSG 433

```

[illegible]



Db 434 VETG--HNRSINRKKYDLAEITAEINQADSLRKETYDALTKLISTRKAESLFHPEIPM 490  
QY 566 TFNTNNKHIIIGYIRNN--ALLAFGNFSEYPTQVTAHTLQAMPFKAHDLIGKTVSLNQD 622  
Db 491 EVLESTAEALFVVKRSSSDAESIILIHNLSE--KEVSYSLDSGVYT--NLTKGSTVTGSDS 545  
QY 623 LTLQPYQVMMLE 634  
Db 546 IKLRGYEFCEWIK 557

RESULT 10  
B85754  
probable polysaccharide hydrolase ycjM [imported] - Escherichia coli (strain O157:H7, su  
C/Species: Escherichia coli  
C/Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C/Accession: B85754  
R.;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A/Reference number: AB5480; MUID:21074935; PMID:11206551  
A/Accession: B85754  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-434 <STO>  
A/Cross-references: GB:AE005174; NID:g12515461; PIDN:AAG56494.1; GSPDB:GN00145; UWGP:Z24  
A/Experimental source: strain O157:H7, substrain EDL933  
C/Genetics:  
A/Gene: ycjM  
C/Superfamily: gtfa protein

Query Match 11.3%; Score 386.5; DB 2; Length 434;  
Best Local Similarity 27.6%; Pred. No. 4.6e-23;  
Matches 129; Conservative 79; Mismatches 201; Indels 59; Gaps 17;  
QY 190 DFIENTSNEHEWAQRCAAGDPLFDNFYYIFPDRMPDQYD-----RTLREIFPDQHP 242  
Db 3 DFVCNHSASKEFKNVYLQQQGFEDFFI-----AVDPQTDLSAVTRPRALPLLTP----- 53  
QY 243 GGFSQLDGD--RWVWTTTNSFQWDLNYSNPVFRAMAGEMFLANLGVLDLRMDAVAFIW 300  
Db 54 ---FQMDNSTRHMTTTFSDQDIDLNRSPLYLLAMVDVLLCYLEKGAEYVRLDAVGFMW 110  
QY 301 KQMGTSCEMLPQAHALIRAFNAVRIAPAVFPEKSEA-IVHPDQVQY-IGQDECQIGYN 358  
Db 111 KERGTSCHLEKTHLIIKLRSIIDDIAPGVITETNVPHKDNIAVFGEGDDEAHMVYQ 170  
QY 359 PLQMALMTLATREVNLLHQALTYRHNLPE-HTAWVNVYRSHDDIGWT-----FADEDA 412  
Db 171 FSLPPLVLAHVQKQNEALCQ-WAQSLSPSGKTFWENFLASHDGIQINPLRGLPESEI 229  
QY 413 AYLIGSYDHRQFLNRFVNRFDGSFARGVFPQYNPSTGDCRVSGTAALVGLAODDPHA 472  
Db 230 LALVEALQOEGALVN--WKNNPDGTRS--PYEINVTYMD-----ALSRESSDEER 276  
QY 473 VDRIKLVSIALSTGGLPLIYLGDEVGTLNDDWSQDSNKSDDSRMAHPRY-NEALYAQ 531  
Db 277 CARFTLAHAITLSPGPVPAIYIQSILGSRND--YAGVEKLGYNRAINRKKYHSEITRE 333  
QY 532 RNDPSTAAGQIYQGLRHMAIVARQSNPRFDGGRVLTENTNNKHIIIGYITNNA---LLAFG 587  
Db 334 LNDEATLRHAAYHELRLITLRSHNEFHPDNNFTIDTVNSSVMRIQSNNAEGNCLTGLF 393  
QY 588 NFSEYPTQVTAHTLQAMPFKAH-DLIGKTVSLNQD LTLQPYQVMMLE 634  
Db 394 NVSKNIQHVNITGLHGRDLISEVDILGNK-----ITLRPWQVMMWIK 434

RESULT 11  
AF1416  
Sucrose phosphorylase homolog lmo2735 [imported] - Listeria monocytogenes (strain EGD-e)  
C/Species: Listeria monocytogenes

C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C/Accession: AF1416  
R.;Glasner, P.; Frangenul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgey, O.; Entian, K.D.; Fsihi, H  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkut, G.; Madueno, E.; Maitournam, A.; M  
ok, C.; Schlueter, T.; Simoes, N.; Tixeret, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland  
A/Title: Comparative genomics of Listeria species.  
A/Reference number: AB1077; MUID:21537279; PMID:11679669  
A/Accession: AF1416  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-560 <GLA>  
A/Cross-references: GB:NC 003210; PIDN:CAD00948.1; PID:g16412235; GSPDB:GN00177  
A/Experimental source: strain EGD-e  
C/Genetics:  
A/Gene: lmo2735

Query Match 10.9%; Score 373; DB 2; Length 560;  
Best Local Similarity 23.2%; Pred. No. 8.2e-22;  
Matches 144; Conservative 132; Mismatches 246; Indels 98; Gaps 26;  
QY 50 KLMNELD----SVYGNNEALLP-MLEMLLAQAWQSYQRNSSLKDIDDIARENNPDWILSN 104  
Db 4 ELVNQIEVKLRKIY--QAAVQPAYLEKMLACA-ENYS--NNTRGSID-----TISE 49  
QY 105 KOYGVACYVD-LFAGD--LKGLKDKIPYFQELGLTYLHMLPLFKCPCKSGDGAYVSSY 160  
Db 50 KNVYLLAYGDSIFEKNKHPLQTLNEFLQYEAQDAITDVHLPLIF--PSTSDGFSVTDY 106  
QY 161 RDVNPALGTIGDLREVIAALHEAGISAVVDFIENHTSNEHEWAQRCAAGDPLFDNFYYIF 220  
Db 107 KQIDEGLGMDWDVQKM-----SENFVYMLDFVANHMSKSSDWFKRFSQNEAPYNGF--F 158  
QY 221 PDRRMPDQYDRTLREIFPDQHPGGSQLEDGRVWVTTFNSFQWDLNYSNPVFRAMAGEM 280  
Db 159 IEKDSQPDYKNVTR--PRISP-LFHXYENGKELWTTFSQDLNVRNIDCLVALTDVL 214  
QY 281 LFLANLGVDLRMDAVAFIWKQMGTSCEMLPQAHALIRAFNAVRIAPAVFPEKSEAIVH 340  
Db 215 LFYASKQATSIRLDALIGLWKTSGTGMHPETHEIISLWRLIDELVNLQIITETNV 274  
QY 341 PDQVQYI--GQDECQIGYNPLQMALMTLATREVNLLHQALTYRHNLPEHTAWVNVYR 398  
Db 275 HEENISYGDGKNEANMVYQPLPLVLAHTFTCHDTKLISKWAKSIQVSDTATYFNFLA 334  
QY 399 SHDDIGW-----TFADEDAAYLGISGYDHRQFLNRFVNRFDGSFARGVFPQYNPSTGDC 453  
Db 335 SHDGIKMRPATGILSDEE-----INSLVQKAVQNGQVSKDNADGTQSVYELNINYE- 388  
QY 454 RVSGTAALVGLAODDPHAVDRIKL--YSIALSTGGLPLIYLGDEVGTLNDDWSQDSN 511  
Db 389 -----ALQNLDEDTTBELVTYKIIAASHILLTLQGVPAIYHSLGSKND--LVGYE 438  
QY 512 KSDDSRMAHPRY-NEALYAQRNDSTAAGQIYQGLRHMAIVARQSNPRF-----D 560  
Db 439 ESGINRRINREKLEKNQVLHELKTD--TYRKTIFTSLKKLVQIRNHTAFSPFATQEIID 496  
QY 561 GGR--LVTFNTNNKHIIIGYIRNNALLAFGNFSEYPTQVTAHTL-QAMPFKAHDLIGK 616  
Db 497 LGPDVFAIKRESEGECTYGI-----NVTSHDISKVAVFGTNLLANQ 540  
QY 617 VSLNQD LTLQPYQVMMLEIA 636  
Db 541 VT--SELELTAYEVVWIKKA 558

RESULT 12  
AF1231  
alpha,alpha-phosphotrehalase homolog lmo1254 [imported] - Listeria monocytogenes (stra  
C/Species: Listeria monocytogenes  
C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001



